STIC-Biot ch/Ch mLib 92268

From: Li, Ruixiang
Sent: Wednesday, April 23, 2003 6:16 PM
To: STIC-Biotech/ChemLib
Subject: Sequence search of Application NO: 09/898,586

Please do a standard search on SEQ ID NO: 23 against commercial nucleic acid databases.

Thank you very much!

Ruixiang Li
GAU 1646
CM1 10E18
Mail Box 10D19
306-0282

Point of Contact:
Toby Port
Technical Info. Specialist
CM1 6A04
703-308-39[illegible]

	TYPE OF SEARCH:	VENDOR/COST (where applic.)
Searcher:_____	NA Sequences:_____	STN:_____
Phone:_____	AA Sequences:_____	DIALOG:_____
Location:_____	Structures:_____	Questel/Orbit:_____
Date Picked Up: 4/24	Bibliographic:_____	DRLink:_____
Date Completed: 5/2	Litigation:_____	Lexis/Nexis:_____
Searcher Prep/Review:_____	Full text:_____	Sequence Sys.:_____
Clerical:_____	Patent Family:_____	WWW/Internet:_____
Online time:_____	Other:_____	Other (specify):_____

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 03:54:25 ; Search time 279 Seconds
(without alignments)
8184.682 Million cell updates/sec

Title: US-09-898-586-23
Perfect score: 1014
Sequence: 1 taaacacttctcctaaacca.....tatggcattgtgactgacaa 1014

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	991	97.7	1014	22	AAS09956	DNA encoding human
	2	987.4	97.4	1012	22	AAS09955	DNA encoding human
	3	983.6	97.0	1040	22	AAS09946	DNA encoding human
c	4	942	92.9	1315	22	ABA09073	Human olfactory re
	5	936	92.3	974	22	AAF58613	Human RECAP polynu
	6	905.4	89.3	930	22	AAH32366	Human olfactory re
	7	903.6	89.1	933	24	ABK65141	cDNA encoding huma
	8	900.6	88.8	930	22	AAH31617	Human olfactory re
	9	900.6	88.8	930	22	AAH31648	Human olfactory re
c	10	545	53.7	1957	22	ABA45207	Human breast cell
c	11	545	53.7	1957	22	ABA55696	Human foetal liver
c	12	545	53.7	1957	22	AAI13969	Probe #3902 for ge
c	13	545	53.7	1957	22	AAI35352	Probe #4038 used t
c	14	545	53.7	1957	24	ABS03949	Human genome-deriv
	15	538.2	53.1	2282	22	AAS08660	Human cDNA encodin
	16	536.8	52.9	940	24	ABN89130	Human GPCR10 nucle
	17	536.6	52.9	1788	22	AAS08655	Human cDNA encodin
	18	536.6	52.9	2735	22	AAS08654	Human cDNA encodin
	19	535.4	52.8	930	22	AAH32363	Human olfactory re
c	20	535.2	52.8	933	22	ABA50343	Human breast cell
c	21	535.2	52.8	933	22	ABA68294	Human foetal liver
c	22	535.2	52.8	933	22	AAI23185	Probe #13118 for g
c	23	535.2	52.8	933	22	AAI48508	Probe #17194 used
c	24	535.2	52.8	933	24	ABS16484	Human genome-deriv
	25	519.8	51.3	813	24	ABK65147	DNA encoding human
	26	518.4	51.1	1792	22	AAF58609	Human RECAP polynu
	27	517.8	51.1	993	22	AAH32365	Human olfactory re
	28	511.4	50.4	1782	22	AAS08657	Human cDNA encodin
	29	481.2	47.5	1713	17	AAT33900	Human G-protein co
	30	481.2	47.5	1713	21	AAA70768	Human G-protein co
	31	481.2	47.5	1713	21	AAZ60116	Human G-protein co
	32	429.6	42.4	576	22	AAH32210	Human olfactory re
	33	415	40.9	952	24	ABQ88400	Human G protein co
	34	415	40.9	963	24	ABQ88399	Human G protein co
	35	414	40.8	933	22	AAS42451	Human cDNA encodin
	36	414	40.8	933	24	ABK37737	DNA encoding G-cou
	37	409.2	40.4	942	22	AAH32170	Human olfactory re
	38	407.2	40.2	937	22	AAH32345	Human olfactory re
	39	405.6	40.0	1411	24	ABK65148	DNA encoding human
	40	402.2	39.7	930	22	AAS42444	Human cDNA encodin
	41	402.2	39.7	930	24	ABK37730	DNA encoding G-cou
	42	398.2	39.3	988	22	AAH32061	Human olfactory re
	43	390	38.5	933	22	AAS42450	Human cDNA encodin
	44	390	38.5	933	24	ABK37736	DNA encoding G-cou
	45	389.8	38.4	938	24	ABN89129	Human GPCR9 nucleo

ALIGNMENTS

RESULT 1
AAS09956
ID AAS09956 standard; DNA; 1014 BP.
XX
AC AAS09956;
XX
DT 24-OCT-2001 (first entry)
XX
DE DNA encoding human odorant receptor (OR)-like protein, NOV12.
XX
KW Odorant receptor; human; OR; NOV12; therapeutic; neuro-olfactory system;
KW trauma; surgery; neoplastic disorder; gene therapy; adenocarcinoma;
KW lymphoma; prostate cancer; uterus cancer; immune response; asthma; AIDS;
KW acquired immunodeficiency syndrome; Crohn's disease; multiple sclerosis;
KW Albright hereditary osteodystrophy; diagnostic; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..54
FT /*tag= a
FT CDS 55..987
FT /*tag= b
FT /product= "Odorant receptor-like protein, NOV12"
FT 3'UTR 988..1014
FT /*tag= c
XX
PN WO200151632-A2.
XX
PD 19-JUL-2001.
XX
PF 16-JAN-2001; 2001WO-US01513.

```
XX
PR   13-JAN-2000; 2000US-0175989.
PR   14-JAN-2000; 2000US-0176134.
PR   25-JAN-2000; 2000US-0177839.
PR   26-JAN-2000; 2000US-0178191.
PR   26-JAN-2000; 2000US-0178227.
PR   14-JUL-2000; 2000US-0218324.
PR   24-JUL-2000; 2000US-0220253.
PR   25-JUL-2000; 2000US-0220590.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Padigaru M,  Prayaga SK,  Taupier RJ,  Mishra V,  Tchernev VT;
PI   Spytek KA,  Li L;
XX
DR   WPI; 2001-451859/48.
DR   P-PSDB; AAU05142.
XX
PT   New NOVX polypeptides and polynucleotides, useful for treating or
PT   preventing disorders of the neuro-olfactory system, cancer and multiple
PT   sclerosis -
XX
PS   Claim 9; Page 55; 141pp; English.
XX
CC   The sequence represents the coding sequence of human odorant receptor
CC   (OR)-like protein, NOV12. The NOV12 polypeptide, nucleic acid and
CC   antibody are useful as therapeutics, particularly in the manufacture of a
CC   medicament for treating a syndrome associated with a human disease, which
CC   includes a pathology associated with NOV12 polypeptide. The NOV12 nucleic
CC   acid and polypeptide are especially useful in therapeutic or
CC   prophylactic applications for disorders of the neuro-olfactory system,
CC   e.g. those induced by trauma, surgery and/or neoplastic disorders. The
CC   DNA encoding the protein is useful in gene therapy for treating the
CC   above conditions. Furthermore, the nucleic acids and polypeptides are
CC   useful in treating adenocarcinoma, lymphoma, prostate cancer, uterus
CC   cancer, immune response, acquired immunodeficiency syndrome (AIDS),
CC   asthma, Crohn's disease, multiple sclerosis or Albright hereditary
CC   osteodystrophy. These are also useful in developing powerful assay
CC   system for functional analysis of various human disorders, as well as
CC   in diagnostic applications.
XX
SQ   Sequence 1014 BP; 219 A; 281 C; 216 G; 298 T; 0 other;

  Query Match             97.7%;  Score 991;  DB 22;  Length 1014;
  Best Local Similarity   99.8%;  Pred. No. 4.3e-287;
  Matches 1013;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    2;

Qy       1 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGG 60

Qy      61 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 120

Qy     121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 180

Qy     181 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTCTTCCTC 239
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     181 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTCTTCCTC 240

Qy     240 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTG 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTG 300

Qy     300 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 360

Qy     360 TTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 420

Qy     420 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 480

Qy     480 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 540

Qy     540 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTTGTGAAATCTTGGC 599
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     541 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAATCTTGGC 599

Qy     600 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 659

Qy     660 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     660 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 719

Qy     720 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTC 779
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     720 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTC 779

Qy     780 CCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 839
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     780 CCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 839

Qy     840 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     840 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 899

Qy     900 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     900 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 959

Qy     960 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACAA 1014
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     960 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACAA 1014


RESULT 2
AAS09955
ID   AAS09955 standard; DNA; 1012 BP.
XX
AC   AAS09955;
XX
DT   24-OCT-2001  (first entry)
XX
DE   DNA encoding human odorant receptor (OR)-like protein, NOV11.
XX
KW   Odorant receptor; human; OR; NOV11; therapeutic; neuro-olfactory system;
KW   trauma; surgery; neoplastic disorder; gene therapy; adenocarcinoma;
KW   lymphoma; prostate cancer; uterus cancer; immune response; asthma; AIDS;
KW   acquired immunodeficiency syndrome; Crohn's disease; multiple sclerosis;
KW   Albright hereditary osteodystrophy; diagnostic; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   5'UTR            1..53
FT                    /*tag=  a
FT   CDS              54..986
FT                    /*tag=  b
FT                    /product= "Odorant receptor-like protein, NOV11"
FT   3'UTR            987..1012
FT                    /*tag=  c
XX
PN   WO200151632-A2.
XX
PD   19-JUL-2001.
```

```
XX
PF   16-JAN-2001; 2001WO-US01513.
XX
PR   13-JAN-2000; 2000US-0175989.
PR   14-JAN-2000; 2000US-0176134.
PR   25-JAN-2000; 2000US-0177839.
PR   26-JAN-2000; 2000US-0178191.
PR   26-JAN-2000; 2000US-0178227.
PR   14-JUL-2000; 2000US-0218324.
PR   24-JUL-2000; 2000US-0220253.
PR   25-JUL-2000; 2000US-0220590.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Padigaru M, Prayaga SK, Taupier RJ, Mishra V, Tchernev VT;
PI   Spytek KA, Li L;
XX
DR   WPI; 2001-451859/48.
DR   P-PSDB; AAU05141.
XX
PT   New NOVX polypeptides and polynucleotides, useful for treating or
PT   preventing disorders of the neuro-olfactory system, cancer and multiple
PT   sclerosis -
XX
PS   Claim 9; Page 52; 141pp; English.
XX
CC   The sequence represents the coding sequence of human odorant receptor
CC   (OR)-like protein, NOV11. The NOV11 polypeptide, nucleic acid and
CC   antibody are useful as therapeutics, particularly in the manufacture of a
CC   medicament for treating a syndrome associated with a human disease, which
CC   includes a pathology associated with NOV11 polypeptide. The NOV11 nucleic
CC   acid and polypeptide are especially useful in therapeutic or
CC   prophylactic applications for disorders of the neuro-olfactory system,
CC   e.g. those induced by trauma, surgery and/or neoplastic disorders. The
CC   DNA encoding the protein is useful in gene therapy for treating the
CC   above conditions. Furthermore, the nucleic acids and polypeptides are
CC   useful in treating adenocarcinoma, lymphoma, prostate cancer, uterus
CC   cancer, immune response, acquired immunodeficiency syndrome (AIDS),
CC   asthma, Crohn's disease, multiple sclerosis or Albright hereditary
CC   osteodystrophy. These are also useful in developing powerful assay
CC   system for functional analysis of various human disorders, as well as
CC   in diagnostic applications.
XX
SQ   Sequence 1012 BP; 219 A; 281 C; 215 G; 297 T; 0 other;

  Query Match             97.4%;  Score 987.4;  DB 22;  Length 1012;
  Best Local Similarity   99.7%;  Pred. No. 5.2e-286;
  Matches 1010;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    2;

Qy       2 AAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGGG 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1 AAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGAG 60

Qy      62 ACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATTC 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATTC 120

Qy     122 AGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGGA 181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGGA 180

Qy     182 CCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTCTTCCTCT 240
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     181 CCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTCTTCCTCT 240

Qy     241 CACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGA 300

Qy     301 ACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTGT 360

Qy     361 TTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTACG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTACG 420

Qy     421 TGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACCC 480

Qy     481 TCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTC 540

Qy     541 TACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAAATCTTGGCT 600
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     541 TACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAATCTTGGCT 599

Qy     601 GTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGCA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 GTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGCA 659

Qy     661 ATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTGT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     660 ATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTGT 719

Qy     721 GCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTCC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     720 GCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTCC 779

Qy     781 CACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     780 CACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAGA 839

Qy     841 TATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCCC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     840 TATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCCC 899

Qy     901 ATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAGA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     900 ATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAGA 959

Qy     961 GTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1013
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     960 GTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1012


RESULT 3
AAS09946
ID   AAS09946 standard; DNA; 1040 BP.
XX
AC   AAS09946;
XX
DT   24-OCT-2001  (first entry)
XX
DE   DNA encoding human odorant receptor (OR)-like protein, NOV2.
XX
KW   Odorant receptor; human; OR; NOV2; therapeutic; neuro-olfactory system;
KW   trauma; surgery; neoplastic disorder; gene therapy; adenocarcinoma;
KW   lymphoma; prostate cancer; uterus cancer; immune response; asthma; AIDS;
KW   acquired immunodeficiency syndrome; Crohn's disease; multiple sclerosis;
KW   Albright hereditary osteodystrophy; diagnostic; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..81
FT                   /*tag= a
FT   CDS             82..1014
FT                   /*tag= b
FT                   /product= "Odorant receptor-like protein, NOV2"
FT   3'UTR           1015..1040
FT                   /*tag= c
XX
PN   WO200151632-A2.
```

XX
PD 19-JUL-2001.
XX
PF 16-JAN-2001; 2001WO-US01513.
XX
PR 13-JAN-2000; 2000US-0175989.
PR 14-JAN-2000; 2000US-0176134.
PR 25-JAN-2000; 2000US-0177839.
PR 26-JAN-2000; 2000US-0178191.
PR 26-JAN-2000; 2000US-0178227.
PR 14-JUL-2000; 2000US-0218324.
PR 24-JUL-2000; 2000US-0220253.
PR 25-JUL-2000; 2000US-0220590.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Padigaru M, Prayaga SK, Taupier RJ, Mishra V, Tchernev VT;
PI Spytek KA, Li L;
XX
DR WPI; 2001-451859/48.
DR P-PSDB; AAU05132.
XX
PT New NOVX polypeptides and polynucleotides, useful for treating or
PT preventing disorders of the neuro-olfactory system, cancer and multiple
PT sclerosis -
XX
PS Claim 9; Page 15; 141pp; English.
XX
CC The sequence represents the coding sequence of human odorant receptor
CC (OR)-like protein, NOV2. The NOV2 polypeptide, nucleic acid and antibody
CC are useful as therapeutics, particularly in the manufacture of a
CC medicament for treating a syndrome associated with a human disease, which
CC includes a pathology associated with NOV2 polypeptide. The NOVX nucleic
CC acid and polypeptide are especially useful in therapeutic or
CC prophylactic applications for disorders of the neuro-olfactory system,
CC e.g. those induced by trauma, surgery and/or neoplastic disorders. The
CC DNA encoding the protein is useful in gene therapy for treating the
CC above conditions. Furthermore, the nucleic acids and polypeptides are
CC useful in treating adenocarcinoma, lymphoma, prostate cancer, uterus
CC cancer, immune response, acquired immunodeficiency syndrome (AIDS),
CC asthma, Crohn's disease, multiple sclerosis or Albright hereditary
CC osteodystrophy. These are also useful in developing powerful assay
CC system for functional analysis of various human disorders, as well as
CC in diagnostic applications.
XX
SQ Sequence 1040 BP; 229 A; 285 C; 221 G; 305 T; 0 other;

Query Match 97.0%; Score 983.6; DB 22; Length 1040;
Best Local Similarity 99.4%; Pred. No. 7.3e-285;
Matches 1008; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

```
Qy       1 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      28 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGA 87

Qy      61 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 120
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db      88 GACAATATAACATCCATCAGAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 147

Qy     121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     148 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 207

Qy     181 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTCTTCCTC 239
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     208 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTCTTCCTC 267

Qy     240 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCCGGATGCTGGTG 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     268 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTG 327

Qy     300 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     328 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 387

Qy     360 TTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     388 TTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 447

Qy     420 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     448 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 507

Qy     480 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     508 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 567

Qy     540 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAAATCTTGGC 599
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     568 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAATCTTGGC 626

Qy     600 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     627 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 686

Qy     660 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     687 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 746

Qy     720 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTC 779
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     747 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCCGCACCTGCTTCTC 806

Qy     780 CCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 839
          ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db     807 CCACCTCTGTGTGATTGGACTCGTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 866

Qy     840 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     867 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 926

Qy     900 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     927 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 986

Qy     960 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1013
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     987 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1040
```

RESULT 4
ABA09073/c
ID ABA09073 standard; cDNA; 1315 BP.
XX
AC ABA09073;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human olfactory receptor homologue-encoding cDNA, SEQ ID NO:849.
XX
KW *Human;* cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell *disorder; asthma; arthritis;*
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW *bone disorder; osteoporosis;* vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; *antibacterial;*
KW antifungal; vulnerary; antiulcer; ss.
XX
OS *Homo sapiens.*
XX

PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03800.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-457740/49.
DR P-PSDB; ABB11829.
XX
PT Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX
PS Claim 1; Page 753; 1963pp; English.
XX
CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g.,myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention.
XX
SQ Sequence 1315 BP; 380 A; 281 C; 360 G; 294 T; 0 other;

Query Match 92.9%; Score 942; DB 22; Length 1315;
Best Local Similarity 99.6%; Pred. No. 2.6e-272;
Matches 986; Conservative 0; Mismatches 0; Indels 4; Gaps 4;

```
Qy       1 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     989 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGG 930

Qy      61 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     929 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 870

Qy     121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     869 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 810

Qy     181 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTCTTCCTC 239
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db     809 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTCTTCCTC 750

Qy     240 TCACACCT-GGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCCGGATGCTGGT 298
          |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     749 TCACACCTGGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCCGGATGCTGGT 690

Qy     299 GAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCT 358
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     689 GAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCT 630

Qy     359 GTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTA 418
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     629 GTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTA 570

Qy     419 CGT-GGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCA 477
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     569 CGTGGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCA 510

Qy     478 CCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTAC 537
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     509 CCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTAC 450

Qy     538 TTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAAATCTTG 597
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db     449 TTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAATCTTG 391

Qy     598 GCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGA 657
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     390 GCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGA 331

Qy     658 GCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTC 717
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     330 GCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTC 271

Qy     718 TGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTC 777
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     270 TGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTC 211

Qy     778 TCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCC 837
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     210 TCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCC 151

Qy     838 AGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAAT 897
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     150 AGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAAT 91

Qy     898 CCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAG 957
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      90 CCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAG 31

Qy     958 AGAGTGCTGGGAGTAGAAAGGGCTTTATGA 987
          |||||||||||||||||||||||||||||
Db      30 AGAGTGCTGGGAGTAGAAAGGGCTTTATGA 1
```

RESULT 5
AAF58613
ID AAF58613 standard; cDNA; 974 BP.
XX
AC AAF58613;
XX
DT 24-APR-2001 (first entry)
XX
DE Human RECAP polynucleotide, SEQ ID NO: 41.
XX
KW Human; RECAP; receptors and associated proteins; cerebroprotective;

```
KW   nootropic; neuroprotective; anticonvulsant; antiparkinsonian; anti-HIV;
KW   antidiabetic; immunostimulant; immunomodulator; antiinflammatory;
KW   antithyroid; immunosuppressive; nephrotropic; antigout; thyromimetic;
KW   cytostatic; antibacterial; virucide; fungicide; protozoacide;
KW   antiarteriosclerotic; hepatotropic; gene therapy; infection; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200107612-A2.
XX
PD   01-FEB-2001.
XX
PF   21-JUL-2000; 2000WO-US20035.
XX
PR   21-JUL-1999;   99US-0145232.
PR   07-OCT-1999;   99US-0158578.
PR   12-NOV-1999;   99US-0165192.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Au-Young J, Bandman O, Tang YT, Yue H, Azimzai Y, Burford N;
PI   Baughn MR, Lu DAM, Hillman JL, Patterson C, Lal P;
XX
DR   WPI; 2001-168554/17.
DR   P-PSDB; AAB68889.
XX
PT   Novel receptors and associated proteins for diagnosis and treatment of
PT   neurological disorders, immunological disorders including autoimmune/
PT   inflammatory disorders and cell proliferative disorders such as cancer
PT   -
XX
PS   Claim 5; Page 125-126; 128pp; English.
XX
CC   The present sequence encodes a human RECAP (receptors and associated
CC   proteins) polypeptide. RECAP polynucleotides and polypeptides are useful
CC   in the diagnosis, treatment and prevention of neurological disorders
CC   such as stroke, Alzheimer's disease, Pick's disease, Huntington's
CC   disease, dementia, Parkinson's disease, Downs's syndrome, amyotrophic
CC   lateral sclerosis, multiple sclerosis, bacterial and viral meningitis,
CC   CJD (Creutzfeldt-Jakob disease), GSS (Gerstmann -Straussler-Scheinker
CC   syndrome); immunological disorders, including autoimmune/inflammatory
CC   disorders such as AIDS, DiGeorge's syndrome, severe combined
CC   immunodeficiency disease (SCID), Chediak-Higashi syndrome, Cushing's
CC   disease, Addison's disease, autoimmune thyroiditis, Crohn's disease,
CC   diabetes mellitus, Good pasture's syndrome, gout, Grave's diseases,
CC   Hashimoto's thyroiditis, Sjogren's syndrome, Werner's syndrome, viral,
CC   bacterial, fungal, parasitic, protozoal, and helminthic infections; and
CC   cell proliferation disorders such as arteriosclerosis, atherosclerosis,
CC   cirrhosis, hepatitis and cancer.
XX
SQ   Sequence 974 BP; 206 A; 270 C; 214 G; 284 T; 0 other;

  Query Match             92.3%;  Score 936;  DB 22;  Length 974;
  Best Local Similarity   99.8%;  Pred. No. 1.4e-270;
  Matches  958;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    2;

Qy      55 ATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCA 114
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCA 60

Qy     115 AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGG 174
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGG 120

Qy     175 AACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTC 233
           ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db     121 AACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTC 180

Qy     234 TTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATG 293
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATG 240

Qy     294 CTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 353
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 300

Qy     354 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGAT 413
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGAT 360

Qy     414 CTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 473
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 420

Qy     474 ATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTG 533
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTG 480

Qy     534 TTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAAAT 593
           |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     481 TTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAAT 539

Qy     594 CTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGC 653
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     540 CTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGC 599

Qy     654 CGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCAT 713
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 CGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCAT 659

Qy     714 CCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTG 773
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     660 CCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTG 719

Qy     774 CTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGG 833
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     720 CTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGG 779

Qy     834 ACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTT 893
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     780 ACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTT 839

Qy     894 TAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTT 953
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     840 TAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTT 899

Qy     954 GAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1013
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     900 GAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 959


RESULT 6
AAH32366
ID   AAH32366 standard; DNA; 930 BP.
XX
AC   AAH32366;
XX
DT   30-JUL-2001  (first entry)
XX
DE   Human olfactory receptor polynucleotide, SEQ ID NO: 939.
XX
KW   Human; olfactory receptor; OR; primary scent determination;
KW   secondary scent determination; polypeptide library; odour receptor;
KW   scent profile; scent fingerprint; scent representation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200127158-A2.
XX
PD   19-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US27582.
XX
PR   08-OCT-1999;   99US-0158615.
PR   24-FEB-2000; 2000US-0184809.
XX
```

```
PA   (DIGI-) DIGISCENTS.
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR   WPI; 2001-290713/30.
XX
PT   New polynucleotides which encode polypeptides involved in olfactory
PT   sensation for identifying olfactory agonists and antagonists -
XX
PS   Claim 8; Page 554; 1857pp; English.
XX
CC   The present sequence is one of a number of isolated polynucleotides
CC   which encode polypeptides involved in olfactory sensation. The
CC   polynucleotides can be used in screening for olfactory agonists and
CC   antagonists. The methods allow for the determination of primary
CC   scents and the identification of the odour receptors used to detect
CC   these primary scents. The methods also enable determination of
CC   secondary scents and the identification of combinations of odour
CC   receptors that are involved in detecting such secondary scents.
CC   This enables the construction of a scent representation (also called
CC   a scent fingerprint or scent profile), which may be used to re-create
CC   and edit scents. Libraries of olfactory receptors are useful for
CC   determining the interaction pattern of a composition with the receptors,
CC   and can be used for determining differences in the olfactory faculties
CC   of different individuals.
XX
SQ   Sequence 930 BP; 192 A; 266 C; 201 G; 271 T; 0 other;

  Query Match             89.3%;  Score 905.4;  DB 22;  Length 930;
  Best Local Similarity   99.7%;  Pred. No. 2.1e-261;
  Matches  928;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    2;

Qy      55 ATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCA 114
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCA 60

Qy     115 AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGG 174
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGG 120

Qy     175 AACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTC 233
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db     121 AACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTC 180

Qy     234 TTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATG 293
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATG 240

Qy     294 CTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 353
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 300

Qy     354 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGAT 413
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGAT 360

Qy     414 CTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 473
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 420

Qy     474 ATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTG 533
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTG 480

Qy     534 TTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAAAT 593
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     481 TTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAAT 539

Qy     594 CTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGC 653
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     540 CTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGC 599

Qy     654 CGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCAT 713
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 CGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCAT 659

Qy     714 CCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTG 773
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     660 CCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTG 719

Qy     774 CTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGG 833
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     720 CTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGG 779

Qy     834 ACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTT 893
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     780 ACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCCGCTGTTTCACAGCCTCTT 839

Qy     894 TAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTT 953
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     840 TAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTT 899

Qy     954 GAAGAGAGTGCTGGGAGTAGAAAGGGCTTTA 984
           |||||||||||||||||||||||||||||||
Db     900 GAAGAGAGTGCTGGGAGTAGAAAGGGCTTTA 930


RESULT 7
ABK65141
ID   ABK65141 standard; cDNA; 933 BP.
XX
AC   ABK65141;
XX
DT   02-JUL-2002  (first entry)
XX
DE   cDNA encoding human PHOR1-F5D6.
XX
KW   Human; PHOR1-A11; PHOR1-F5D6; prostate cancer; cytostatic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200214501-A2.
XX
PD   21-FEB-2002.
XX
PF   17-AUG-2001; 2001WO-US25862.
XX
PR   17-AUG-2000; 2000US-226241P.
XX
PA   (AGEN-) AGENSYS INC.
XX
PI   Hubert RS, Raitano AB, Faris M, Challita-Eid PM, Ge W;
PI   Jakobovits A;
XX
DR   WPI; 2002-269193/31.
DR   P-PSDB; AAU91543.
XX
PT   Monitoring PHOR1-A11/PHOR1-F5D6 gene products for monitoring presence
PT   of cancer in subject, by determining status of PHOR1-A11/PHOR1-F5D6
PT   gene products in tissue sample from subject and comparing it to normal
PT   sample -
XX
PS   Claim 43; Fig 2b; 250pp; English.
XX
CC   The present invention relates to the isolation of novel human genes
CC   designated PHOR1-A11 and PHOR1-F5D6 and their encoded proteins. The
CC   gene encoding PHOR1-A11 maps to chromosome 1q23, and the gene encoding
CC   PHOR1-F5D6 maps to chromosome 7q33-q35. The PHOR1-A11 and PHOR1-F5D6
CC   polynucleotide and polypeptide sequences are useful in diagnostic and
CC   therapeutic methods, and compositions for various cancers such as
CC   prostate cancer. The sequences are useful for inhibiting the growth of
CC   cancer cells that express PHOR1-A11 or PHOR1-F5D6 and for treating
CC   cancer. The PHOR1-A11 or PHOR1-F5D6 polypeptide or a fragment thereof
CC   can be used to elicit an immune response. The present sequence encodes
CC   human PHOR1-F5D6.
```

```
XX
SQ   Sequence 933 BP; 194 A; 265 C; 203 G; 271 T; 0 other;

  Query Match             89.1%;  Score 903.6;  DB 24;  Length 933;
  Best Local Similarity   99.4%;  Pred. No. 7.4e-261;
  Matches  928;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy      55 ATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCA 114
           ||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db       1 ATGGGAGACAATATAACATCCATCAGAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCA 60

Qy     115 AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGG 174
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGG 120

Qy     175 AACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTC 233
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db     121 AACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTC 180

Qy     234 TTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATG 293
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATG 240

Qy     294 CTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 353
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTGCGGGCCGCATGATGCAGACC 300

Qy     354 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGAT 413
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGAT 360

Qy     414 CTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 473
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 420

Qy     474 ATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTG 533
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTG 480

Qy     534 TTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTTGTGAAAT 593
           |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     481 TTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAAT 539

Qy     594 CTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGC 653
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     540 CTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGC 599

Qy     654 CGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCAT 713
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 CGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCAT 659

Qy     714 CCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTG 773
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     660 CCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCCGCACCTG 719

Qy     774 CTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGG 833
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db     720 CTTCTCCCACCTCTGTGTGATTGGACTCGTTTATGGCACAGCCATTATCATGTATGTTGG 779

Qy     834 ACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTT 893
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     780 ACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTT 839

Qy     894 TAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTT 953
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     840 TAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTT 899

Qy     954 GAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGA 987
           |||||||||||||||||||||||||||||||||
Db     900 GAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGA 933


RESULT 8
```

```
AAH31617
ID   AAH31617 standard; DNA; 930 BP.
XX
AC   AAH31617;
XX
DT   30-JUL-2001  (first entry)
XX
DE   Human olfactory receptor polynucleotide, SEQ ID NO: 190.
XX
KW   Human; olfactory receptor; OR; primary scent determination;
KW   secondary scent determination; polypeptide library; odour receptor;
KW   scent profile; scent fingerprint; scent representation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200127158-A2.
XX
PD   19-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US27582.
XX
PR   08-OCT-1999;   99US-0158615.
PR   24-FEB-2000; 2000US-0184809.
XX
PA   (DIGI-) DIGISCENTS.
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Bellenson J,  Smith D,  Lancet D,  Glusman G,  Fuchs T,  Yanai I;
XX
DR   WPI; 2001-290713/30.
XX
PT   New polynucleotides which encode polypeptides involved in olfactory
PT   sensation for identifying olfactory agonists and antagonists -
XX
PS   Claim 8; Page 251; 1857pp; English.
XX
CC   The present sequence is one of a number of isolated polynucleotides
CC   which encode polypeptides involved in olfactory sensation. The
CC   polynucleotides can be used in screening for olfactory agonists and
CC   antagonists. The methods allow for the determination of primary
CC   scents and the identification of the odour receptors used to detect
CC   these primary scents. The methods also enable determination of
CC   secondary scents and the identification of combinations of odour
CC   receptors that are involved in detecting such secondary scents.
CC   This enables the construction of a scent representation (also called
CC   a scent fingerprint or scent profile), which may be used to re-create
CC   and edit scents. Libraries of olfactory receptors are useful for
CC   determining the interaction pattern of a composition with the receptors,
CC   and can be used for determining differences in the olfactory faculties
CC   of different individuals.
XX
SQ   Sequence 930 BP; 193 A; 265 C; 202 G; 270 T; 0 other;

  Query Match             88.8%;  Score 900.6;  DB 22;  Length 930;
  Best Local Similarity   99.4%;  Pred. No. 5.9e-260;
  Matches  925;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy      55 ATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCA 114
           ||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db       1 ATGGGAGACAATATAACATCCATCAGAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCA 60

Qy     115 AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGG 174
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGG 120

Qy     175 AACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTC 233
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db     121 AACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTC 180

Qy     234 TTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATG 293
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATG 240
```

```
Qy    294 CTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 300

Qy    354 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGAT 413
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGAT 360

Qy    414 CTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 473
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 420

Qy    474 ATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTG 533
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTG 480

Qy    534 TTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTTGTGAAAT 593
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    481 TTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAAT 539

Qy    594 CTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGC 653
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 CTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGC 599

Qy    654 CGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCAT 713
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 CGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCAT 659

Qy    714 CCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTG 773
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    660 CCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCCGCACCTG 719

Qy    774 CTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGG 833
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    720 CTTCTCCCACCTCTGTGTGATTGGACTCGTTTATGGCACAGCCATTATCATGTATGTTGG 779

Qy    834 ACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTT 893
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    780 ACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTT 839

Qy    894 TAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTT 953
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    840 TAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTT 899

Qy    954 GAAGAGAGTGCTGGGAGTAGAAAGGGCTTTA 984
          |||||||||||||||||||||||||||||||
Db    900 GAAGAGAGTGCTGGGAGTAGAAAGGGCTTTA 930
```

RESULT 9
AAH31648
ID AAH31648 standard; DNA; 930 BP.
XX
AC AAH31648;
XX
DT 30-JUL-2001 (first entry)
XX
DE Human olfactory receptor polynucleotide, SEQ ID NO: 221.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation; ds.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Claim 8; Page 264-265; 1857pp; English.
XX
CC The present sequence is one of a number of isolated polynucleotides
CC which encode polypeptides involved in olfactory sensation. The
CC polynucleotides can be used in screening for olfactory agonists and
CC antagonists. The methods allow for the determination of primary
CC scents and the identification of the odour receptors used to detect
CC these primary scents. The methods also enable determination of
CC secondary scents and the identification of combinations of odour
CC receptors that are involved in detecting such secondary scents.
CC This enables the construction of a scent representation (also called
CC a scent fingerprint or scent profile), which may be used to re-create
CC and edit scents. Libraries of olfactory receptors are useful for
CC determining the interaction pattern of a composition with the receptors,
CC and can be used for determining differences in the olfactory faculties
CC of different individuals.
XX
SQ Sequence 930 BP; 193 A; 265 C; 202 G; 270 T; 0 other;

Query Match 88.8%; Score 900.6; DB 22; Length 930;
Best Local Similarity 99.4%; Pred. No. 5.9e-260;
Matches 925; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

```
Qy     55 ATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCA 114
          |||||| |||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      1 ATGGGAGACAATATAACATCCATCAGAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCA 60

Qy    115 AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGG 174
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGG 120

Qy    175 AACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTC 233
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db    121 AACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTC 180

Qy    234 TTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATG 293
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATG 240

Qy    294 CTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 300

Qy    354 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGAT 413
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGAT 360

Qy    414 CTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 473
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 420

Qy    474 ATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTG 533
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTG 480

Qy    534 TTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTTGTGAAAT 593
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    481 TTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAAT 539

Qy    594 CTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGC 653
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 CTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGC 599
```

```
Qy    654 CGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCAT 713
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 CGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCAT 659

Qy    714 CCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTG 773
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    660 CCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCCGCACCTG 719

Qy    774 CTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGG 833
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    720 CTTCTCCCACCTCTGTGTGATTGGACTCGTTTATGGCACAGCCATTATCATGTATGTTGG 779

Qy    834 ACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTT 893
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    780 ACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTT 839

Qy    894 TAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTT 953
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    840 TAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTT 899

Qy    954 GAAGAGAGTGCTGGGAGTAGAAAGGGCTTTA 984
          |||||||||||||||||||||||||||||||
Db    900 GAAGAGAGTGCTGGGAGTAGAAAGGGCTTTA 930
```

RESULT 10
ABA45207/c
ID ABA45207 standard; DNA; 1957 BP.
XX
AC ABA45207;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human breast cell single exon nucleic acid probe #3902.
XX
KW Human; microarray; single exon probe; gene expression; breast;
KW disease; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00662.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX
PS Claim 1; SEQ ID NO 3902; 327pp + sequence listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1957 BP; 604 A; 368 C; 495 G; 490 T; 0 other;

```
  Query Match             53.7%;  Score 545;  DB 22;  Length 1957;
  Best Local Similarity   74.6%;  Pred. No. 6e-153;
  Matches  711;  Conservative    0;  Mismatches  240;  Indels    2;  Gaps    2;

Qy      36 TTTCCTCTGTCATAGGGATATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACT 95
           ||| || | ||| ||||| |||||||| ||| ||| ||||||||||||||||||||||
Db    1073 TTTTCTTTTTCACAGGGAAATGGGGGAAAATCAGACAATGGTCACAGAGTTCCTCCTACT 1014

Qy      96 GGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTA 155
           |||||||| | | ||||||||||||||||||||||||||||||||||||||||||||||
Db    1013 GGGATTTCTCCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTA 954

Qy     156 CGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACT 215
           ||||||||||||||||||||||| |||| ||||||||||||||||||||||||||||||
Db     953 TATCTTCACCCTGCTGGGGAACGGGGCCATCCTGGGGCTCATCTCACTGGACTCCAGACT 894

Qy     216 GCACGCCCCC-TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTG 274
           ||| ||||| |||||||||||||||||||||||| |||||||||||||||||| || |
Db     893 CCACACCCCCATGTACTTCTTCCTCTCACACCTGGCTGTCGTCGACATCGCCTACACCCG 834

Qy     275 CAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGC 334
           ||||||||||||| |||||||| ||||||||||||||||||||||||||||||||||||
Db     833 CAACACGGTGCCCCAGATGCTGGCGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGC 774

Qy     335 GGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCT 394
           ||  |||||| |||||||||||| | ||  | |||||   ||  |||||||||| ||
Db     773 TGGTTGCATGACGCAGACCTTTCTCTGTTTGAGTTTTGGACACAGCGAATGTCTCCTGCT 714

Qy     395 GGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCAT 454
           |||| |||||||||| |||| ||||||||||||||||||||| |||||||| | | |||
Db     713 GGTGCTGATGTCCTACGATCGTTACGTGGCCATCTGCCACCCTCTCCGATACTCCGTCAT 654

Qy     455 CATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTT 514
           |||||||||||||||||||||||||||| || || |||||||||||   |||  |||  |
Db     653 CATGACCTGGAGAGTCTGCATCACCCTGGCCGTCACTTCCTGGACGTGTGGCTCCCTCCT 594

Qy     515 ATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTA 574
           |  ||   ||| | ||  |  | |||    | ||||||| | |||| ||  ||||  |
Db     593 GGCTCTGGCCCATGTGGTTCTCATCCTAAGACTGCCCTTCTCTGGGCCTCATGAAATCAA 534

Qy     575 TCACTTTTTTTTGTGAAATCTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCA 634
           ||| || || |||||||||| || |||| |||| || |||||||| || |||    |||
Db     533 CCAC-TTCTTCTGTGAAATCCTGTCTGTCCTCAGGCTGGCCTGTGCTGACACCTGGCTCA 475

Qy     635 ATGAGAACATGGTCTTGGCCGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTG 694
           |  ||    |  |||| || |       | |     |||||||| ||       |    |
Db     474 ACCAGGTGGTCATCTTTGCAGCCTGCGTGTTCTTCCTGGTGGGGCCACCCAGCCTGGTGC 415

Qy     695 TAGTTTCATATATGTGCATCCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGA 754
           | || || ||   |  ||||||    || |||||   |||||| ||  |||| |  |  |
Db     414 TTGTCTCCTACTCGCACATCCTGGCGGCCATCCTGAGGATCCAGTCTGGGGAGGGCCGCA 355

Qy     755 GGAAAGCCTTCTGCACCTGCTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAG 814
           | || ||||||| |||||||| |||||||||||| ||| | |||||||| | |||||  |
Db     354 GAAAGGCCTTCTCCACCTGCTCCTCCCACCTCTGCGTGGTGGGACTCTTCTTTGGCAGTG 295

Qy     815 CCATTATCATGTATGTTGGACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCC 874
           |||| ||||||||  | |  ||||  |   |  | ||  |||||||| | ||     |
```

```
Db    294 CCATCATCATGTACATGGCCCCCAAGTCCCGCCATCCTGAGGAGCAGCAAAAGGTCTTTT 235

Qy    875 TGCTGTTTCACAGCCTCTTTAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACT 934
          | || ||| |||| | || || || | || || ||||| || | || || ||||||
Db    234 TTCTATTTTACAGTTTTTTCAACCCAACACTTAACCCCCTGATTTACAGCCTGAGGAACG 175

Qy    935 CAGAAGTGAAGAATACTTTGAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGA 987
           ||| || ||| | | ||| |||||  |||||    |||||   ||  | |
Db    174 GAGAGGTCAAGGGTGCCCTGAGGAGAGCACTGGGCAAGGAAAGTCATTCCTAA 122


RESULT 11
ABA55696/c
ID   ABA55696 standard; DNA; 1957 BP.
XX
AC   ABA55696;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #4001.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00669.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human fetal liver -
XX
PS   Claim 1; SEQ ID NO 4001; 639pp + sequence listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for
CC   measuring human gene expression in a sample derived from human foetal
CC   liver. The single exon nucleic acid probes may be used for predicting,
CC   measuring and displaying gene expression in samples derived from human
CC   fetal liver. The present sequence is a single exon nucleic acid
CC   probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1957 BP; 604 A; 368 C; 495 G; 490 T; 0 other;

  Query Match             53.7%;  Score 545;  DB 22;  Length 1957;
  Best Local Similarity   74.6%;  Pred. No. 6e-153;
  Matches  711;  Conservative    0;  Mismatches  240;  Indels    2;  Gaps    2;

Qy     36 TTTCCTCTGTCATAGGGATATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACT 95
          ||| || | ||| ||||| |||||||| |||   |||    ||||||||||||||||||
Db   1073 TTTTCTTTTTCACAGGGAAATGGGGGAAAATCAGACAATGGTCACAGAGTTCCTCCTACT 1014

Qy     96 GGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTA 155
          |||||||| | | ||||||||||||||||||||||||||||||||||||||||||||||
Db   1013 GGGATTTCTCCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTA 954

Qy    156 CGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACT 215
          |||||||||||||||||||||||| |||| ||||||||||||||||||||||||||||
Db    953 TATCTTCACCCTGCTGGGGAACGGGGCCATCCTGGGGCTCATCTCACTGGACTCCAGACT 894

Qy    216 GCACGCCCCC-TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTG 274
          ||| ||||| |||||||||||||||||||||||| |||||||||||||||||| || |
Db    893 CCACACCCCCATGTACTTCTTCCTCTCACACCTGGCTGTCGTCGACATCGCCTACACCCG 834

Qy    275 CAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGC 334
          |||||||||||||| |||||||| ||||||||||||||||||||||||||||||||||||
Db    833 CAACACGGTGCCCCAGATGCTGGCGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGC 774

Qy    335 GGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCT 394
          ||  |||||| |||||||||||| | ||  | |||||    ||  ||||||||||| ||
Db    773 TGGTTGCATGACGCAGACCTTTCTCTGTTTGAGTTTTGGACACAGCGAATGTCTCCTGCT 714

Qy    395 GGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCAT 454
          |||| |||||||||| |||| ||||||||||||||||||| |||||||| |  | |||
Db    713 GGTGCTGATGTCCTACGATCGTTACGTGGCCATCTGCCACCCTCTCCGATACTCCGTCAT 654

Qy    455 CATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTT 514
          |||||||||||||||||||||||||| || || ||||||||||   |||  ||| |
Db    653 CATGACCTGGAGAGTCTGCATCACCCTGGCCGTCACTTCCTGGACGTGTGGCTCCCTCCT 594

Qy    515 ATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTA 574
          |  ||   ||| | ||  |  | |||    | |||||||  | |||| ||  |||| |
Db    593 GGCTCTGGCCCATGTGGTTCTCATCCTAAGACTGCCCTTCTCTGGGCCTCATGAAATCAA 534

Qy    575 TCACTTTTTTTTGTGAAATCTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCA 634
          ||| || || |||||||||| || |||| |||| || |||||||| || |||    |||
Db    533 CCAC-TTCTTCTGTGAAATCCTGTCTGTCCTCAGGCTGGCCTGTGCTGACACCTGGCTCA 475

Qy    635 ATGAGAACATGGTCTTGGCCGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTG 694
          |  ||    |  |||| || |       | |      |||||||| ||       |    |
Db    474 ACCAGGTGGTCATCTTTGCAGCCTGCGTGTTCTTCCTGGTGGGGCCACCCAGCCTGGTGC 415

Qy    695 TAGTTTCATATATGTGCATCCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGA 754
          | || || ||   |  |||||    || |||||  ||||||  || |||| |  |  |
Db    414 TTGTCTCCTACTCGCACATCCTGGCGGCCATCCTGAGGATCCAGTCTGGGGAGGGCCGCA 355

Qy    755 GGAAAGCCTTCTGCACCTGCTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAG 814
          | || |||||||  |||||||| |||||||||||| ||| | ||||||| | |||||  |
Db    354 GAAAGGCCTTCTCCACCTGCTCCTCCCACCTCTGCGTGGTGGGACTCTTCTTTGGCAGTG 295

Qy    815 CCATTATCATGTATGTTGGACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCC 874
          |||| ||||||||  | |  ||||  |   |  | ||  |||||||| | ||     |
Db    294 CCATCATCATGTACATGGCCCCCAAGTCCCGCCATCCTGAGGAGCAGCAAAAGGTCTTTT 235

Qy    875 TGCTGTTTCACAGCCTCTTTAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACT 934
          | || ||| |||| | || || || | || || ||||| || | || || ||||||
Db    234 TTCTATTTTACAGTTTTTTCAACCCAACACTTAACCCCCTGATTTACAGCCTGAGGAACG 175

Qy    935 CAGAAGTGAAGAATACTTTGAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGA 987
          ||| || |||  | |  ||| |||||  |||||    |||||   ||  | |
Db    174 GAGAGGTCAAGGGTGCCCTGAGGAGAGCACTGGGCAAGGAAAGTCATTCCTAA 122


RESULT 12
AAI13969/c
ID   AAI13969 standard; DNA; 1957 BP.
XX
AC   AAI13969;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #3902 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
```

```
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 3902; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1957 BP; 604 A; 368 C; 495 G; 490 T; 0 other;

  Query Match             53.7%;  Score 545;  DB 22;  Length 1957;
  Best Local Similarity   74.6%;  Pred. No. 6e-153;
  Matches  711;  Conservative    0;  Mismatches  240;  Indels    2;  Gaps    2;

Qy      36 TTTCCTCTGTCATAGGGATATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACT 95
           ||| || | ||| ||||| |||||||| ||| ||| ||||||||||||||||||||
Db    1073 TTTTCTTTTTCACAGGGAAATGGGGGAAAATCAGACAATGGTCACAGAGTTCCTCCTACT 1014

Qy      96 GGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTA 155
           |||||||| | | ||||||||||||||||||||||||||||||||||||||||||||||
Db    1013 GGGATTTCTCCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTA 954

Qy     156 CGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACT 215
           ||||||||||||||||||||||| |||| ||||||||||||||||||||||||||||||
Db     953 TATCTTCACCCTGCTGGGGAACGGGGCCATCCTGGGGCTCATCTCACTGGACTCCAGACT 894

Qy     216 GCACGCCCCC-TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTG 274
           ||| ||||| |||||||||||||||||||||||||| ||||||||||||||||| || |
Db     893 CCACACCCCCATGTACTTCTTCCTCTCACACCTGGCTGTCGTCGACATCGCCTACACCCG 834

Qy     275 CAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGC 334
           |||||||||||||| |||||||| ||||||||||||||||||||||||||||||||||||
Db     833 CAACACGGTGCCCCAGATGCTGGCGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGC 774

Qy     335 GGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCT 394
           ||  |||||| |||||||||||| | ||  | |||||    ||  ||||||||||| ||
Db     773 TGGTTGCATGACGCAGACCTTTCTCTGTTTGAGTTTTGGACACAGCGAATGTCTCCTGCT 714

Qy     395 GGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCAT 454
           |||| ||||||||||| ||||  ||||||||||||||||||| |||||||| |  | |||
Db     713 GGTGCTGATGTCCTACGATCGTTACGTGGCCATCTGCCACCCTCTCCGATACTCCGTCAT 654

Qy     455 CATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTT 514
           ||||||||||||||||||||||||||| || || ||||||||||   |||   |||  |
Db     653 CATGACCTGGAGAGTCTGCATCACCCTGGCCGTCACTTCCTGGACGTGTGGCTCCCTCCT 594

Qy     515 ATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTA 574
             |  ||   ||| | ||  |  | |||    | ||||||| | |||| ||  |||| |
Db     593 GGCTCTGGCCCATGTGGTTCTCATCCTAAGACTGCCCTTCTCTGGGCCTCATGAAATCAA 534

Qy     575 TCACTTTTTTTTGTGAAATCTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCA 634
           ||| || || |||||||||| || |||| |||| || |||||||| || |||   |||
Db     533 CCAC-TTCTTCTGTGAAATCCTGTCTGTCCTCAGGCTGGCCTGTGCTGACACCTGGCTCA 475

Qy     635 ATGAGAACATGGTCTTGGCCGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTG 694
           |  ||    |  |||| || |      | |     |||||||| ||      |    |
Db     474 ACCAGGTGGTCATCTTTGCAGCCTGCGTGTTCTTCCTGGTGGGGCCACCCAGCCTGGTGC 415

Qy     695 TAGTTTCATATATGTGCATCCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGA 754
           | || || ||   |  ||||||   || |||||  |||||| ||  |||| |  | |
Db     414 TTGTCTCCTACTCGCACATCCTGGCGGCCATCCTGAGGATCCAGTCTGGGGAGGGCCGCA 355

Qy     755 GGAAAGCCTTCTGCACCTGCTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAG 814
           | || |||||| |||||||| |||||||||||| ||| | |||||||| | |||||  |
Db     354 GAAAGGCCTTCTCCACCTGCTCCTCCCACCTCTGCGTGGTGGGACTCTTCTTTGGCAGTG 295

Qy     815 CCATTATCATGTATGTTGGACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCC 874
           |||| ||||||||  | |  |||| |   |  | ||  |||||||| | ||     |
Db     294 CCATCATCATGTACATGGCCCCCAAGTCCCGCCATCCTGAGGAGCAGCAAAAGGTCTTTT 235

Qy     875 TGCTGTTTCACAGCCTCTTTAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACT 934
           | || ||| |||| | || || || |  || || ||||| || |  || || ||||||
Db     234 TTCTATTTTACAGTTTTTTCAACCCAACACTTAACCCCCTGATTTACAGCCTGAGGAACG 175

Qy     935 CAGAAGTGAAGAATACTTTGAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGA 987
           ||| || |||  | |  ||| |||||  |||||    |||||   ||  | |
Db     174 GAGAGGTCAAGGGTGCCCTGAGGAGAGCACTGGGCAAGGAAAGTCATTCCTAA 122


RESULT 13
AAI35352/c
ID   AAI35352 standard; DNA; 1957 BP.
XX
AC   AAI35352;
XX
DT   17-OCT-2001 (first entry)
XX
DE   Probe #4038 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00663.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human placenta -
XX
PS   Claim 25; SEQ ID No 4038; 654pp; English.
```

```
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders.
XX
SQ   Sequence 1957 BP; 604 A; 368 C; 495 G; 490 T; 0 other;

  Query Match             53.7%;  Score 545;  DB 22;  Length 1957;
  Best Local Similarity   74.6%;  Pred. No. 6e-153;
  Matches  711;  Conservative    0;  Mismatches  240;  Indels    2;  Gaps    2;

Qy      36 TTTCCTCTGTCATAGGGATATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACT 95
           ||| || | ||| ||||| |||||||| ||| ||| |||||||||||||||||||||
Db    1073 TTTTCTTTTTCACAGGGAAATGGGGGAAAATCAGACAATGGTCACAGAGTTCCTCCTACT 1014

Qy      96 GGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTA 155
           |||||||| | | ||||||||||||||||||||||||||||||||||||||||||||||
Db    1013 GGGATTTCTCCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTA 954

Qy     156 CGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACT 215
           ||||||||||||||||||||||| |||| ||||||||||||||||||||||||||||||
Db     953 TATCTTCACCCTGCTGGGGAACGGGGCCATCCTGGGGCTCATCTCACTGGACTCCAGACT 894

Qy     216 GCACGCCCCC-TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTG 274
           ||| ||||| |||||||||||||||||||||||||| ||||||||||||||||| || |
Db     893 CCACACCCCCATGTACTTCTTCCTCTCACACCTGGCTGTCGTCGACATCGCCTACACCCG 834

Qy     275 CAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGC 334
           |||||||||||||| |||||||| ||||||||||||||||||||||||||||||||||||
Db     833 CAACACGGTGCCCCAGATGCTGGCGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGC 774

Qy     335 GGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCT 394
           || |||||| |||||||||||| | ||  | |||||    ||  ||||||||||| ||
Db     773 TGGTTGCATGACGCAGACCTTTCTCTGTTTGAGTTTTGGACACAGCGAATGTCTCCTGCT 714

Qy     395 GGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCAT 454
           |||| |||||||||| ||||  |||||||||||||||||||| |||||||| |  | |||
Db     713 GGTGCTGATGTCCTACGATCGTTACGTGGCCATCTGCCACCCTCTCCGATACTCCGTCAT 654

Qy     455 CATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTT 514
           ||||||||||||||||||||||||||| || || ||||||||||   |||  |||  |
Db     653 CATGACCTGGAGAGTCTGCATCACCCTGGCCGTCACTTCCTGGACGTGTGGCTCCCTCCT 594

Qy     515 ATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTA 574
           |  ||   ||| | ||  |  | |||   | |||||| | |||| ||  |||| |
Db     593 GGCTCTGGCCCATGTGGTTCTCATCCTAAGACTGCCCTTCTCTGGGCCTCATGAAATCAA 534

Qy     575 TCACTTTTTTTTGTGAAATCTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCA 634
           ||| || || |||||||| || |||| |||  || |||||||| || |||   |||
Db     533 CCAC-TTCTTCTGTGAAATCCTGTCTGTCCTCAGGCTGGCCTGTGCTGACACCTGGCTCA 475

Qy     635 ATGAGAACATGGTCTTGGCCGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTG 694
           |  ||    |  |||| || |   | |     |||||||| ||     |   |
Db     474 ACCAGGTGGTCATCTTTGCAGCCTGCGTGTTCTTCCTGGTGGGGCCACCCAGCCTGGTGC 415

Qy     695 TAGTTTCATATATGTGCATCCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGA 754
           | || || ||  |  ||||||   || |||||  |||||| ||  |||| |  |  |
Db     414 TTGTCTCCTACTCGCACATCCTGGCGGCCATCCTGAGGATCCAGTCTGGGGAGGGCCGCA 355

Qy     755 GGAAAGCCTTCTGCACCTGCTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAG 814
           | || ||||||| |||||||| |||||||||||| ||| | |||||||| | |||||  |
Db     354 GAAAGGCCTTCTCCACCTGCTCCTCCCACCTCTGCGTGGTGGGACTCTTCTTTGGCAGTG 295

Qy     815 CCATTATCATGTATGTTGGACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCC 874
           |||| ||||||||  | |  |||| |  |  | ||  |||||||| | ||     |
Db     294 CCATCATCATGTACATGGCCCCCAAGTCCCGCCATCCTGAGGAGCAGCAAAAGGTCTTTT 235

Qy     875 TGCTGTTTCACAGCCTCTTTAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACT 934
           | || ||| ||||  | || || || |  || || ||||| || |  || || ||||||
Db     234 TTCTATTTTACAGTTTTTTCAACCCAACACTTAACCCCCTGATTTACAGCCTGAGGAACG 175

Qy     935 CAGAAGTGAAGAATACTTTGAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGA 987
           ||| || |||  | |  ||| |||||  |||||    |||||   ||  | |
Db     174 GAGAGGTCAAGGGTGCCCTGAGGAGAGCACTGGGCAAGGAAAGTCATTCCTAA 122


RESULT 14
ABS03949/c
ID   ABS03949 standard; DNA; 1957 BP.
XX
AC   ABS03949;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe from lung SEQ ID No 3940.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US00665.
XX
PR   04-FEB-2000; 2000US-180312P.
PR   26-MAY-2000; 2000US-207456P.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-234687P.
PR   27-SEP-2000; 2000US-236359P.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples -
XX
PS   Claim 1; SEQ ID No 3940; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of
CC   probes; the novel set of probes which hybridise at high stringency to a
CC   nucleic acid expressed in the human lung; measuring gene expression in a
CC   sample derived from human lung, comprising (a) contacting the array with
CC   a collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of
CC   the array; identifying exons in a eukaryotic genome, comprising
CC   (a) algorithmically predicting at least one exon from genomic sequences
CC   of the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
```

```
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene
CC   expression analysis, and for identifying exons in a gene, particularly
CC   using human lung derived mRNA and for the study of lung diseases
CC   such as asthma, lung cancer, chronic obstructive pulmonary disease
CC   (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC   fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC   Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC   haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomtosis,
CC   pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC   pulmonary dysplasia, primary ciliary dyskinesis, pulmonary hypertension
CC   and hyaline membrane disease. The present sequence is a single exon
CC   probe of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1957 BP; 604 A; 368 C; 495 G; 490 T; 0 other;

  Query Match             53.7%;  Score 545;  DB 24;  Length 1957;
  Best Local Similarity   74.6%;  Pred. No. 6e-153;
  Matches  711;  Conservative    0;  Mismatches  240;  Indels    2;  Gaps    2;

Qy      36 TTTCCTCTGTCATAGGGATATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACT 95
           ||| || | ||| ||||| |||||||| ||| |||    ||||||||||||||||||||
Db    1073 TTTTCTTTTTCACAGGGAAATGGGGGAAAATCAGACAATGGTCACAGAGTTCCTCCTACT 1014

Qy      96 GGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTA 155
           |||||||| | | ||||||||||||||||||||||||||||||||||||||||||||||
Db    1013 GGGATTTCTCCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTA 954

Qy     156 CGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACT 215
           ||||||||||||||||||||||| |||| ||||||||||||||||||||||||||||||
Db     953 TATCTTCACCCTGCTGGGGAACGGGGCCATCCTGGGGCTCATCTCACTGGACTCCAGACT 894

Qy     216 GCACGCCCCC-TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTG 274
           ||| ||||| |||||||||||||||||||||||||| ||||||||||||||||||| || |
Db     893 CCACACCCCCATGTACTTCTTCCTCTCACACCTGGCTGTCGTCGACATCGCCTACACCCG 834

Qy     275 CAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGC 334
           |||||||||||||| |||||||| ||||||||||||||||||||||||||||||||||||
Db     833 CAACACGGTGCCCCAGATGCTGGCGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGC 774

Qy     335 GGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCT 394
           ||  |||||| |||||||||||| | ||  | |||||    ||  ||||||||||| ||
Db     773 TGGTTGCATGACGCAGACCTTTCTCTGTTTGAGTTTTGGACACAGCGAATGTCTCCTGCT 714

Qy     395 GGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCAT 454
           |||| |||||||||| |||| ||||||||||||||||||| |||||||| |   | |||
Db     713 GGTGCTGATGTCCTACGATCGTTACGTGGCCATCTGCCACCCTCTCCGATACTCCGTCAT 654

Qy     455 CATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTT 514
           ||||||||||||||||||||||||||| || || ||||||||||   |||   |||  |
Db     653 CATGACCTGGAGAGTCTGCATCACCCTGGCCGTCACTTCCTGGACGTGTGGCTCCCTCCT 594

Qy     515 ATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTA 574
           |  ||   ||| | ||  |  | |||    | ||||||| | |||| ||  ||||  |
Db     593 GGCTCTGGCCCATGTGGTTCTCATCCTAAGACTGCCCTTCTCTGGGCCTCATGAAATCAA 534

Qy     575 TCACTTTTTTTTGTGAAATCTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCA 634
           ||| || || |||||||| || |||| ||||  || |||||||| || |||    |||
Db     533 CCAC-TTCTTCTGTGAAATCCTGTCTGTCCTCAGGCTGGCCTGTGCTGACACCTGGCTCA 475

Qy     635 ATGAGAACATGGTCTTGGCCGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTG 694
           |  ||    |  |||| || |       | |     |||||||| ||       |    |
Db     474 ACCAGGTGGTCATCTTTGCAGCCTGCGTGTTCTTCCTGGTGGGGCCACCCAGCCTGGTGC 415

Qy     695 TAGTTTCATATATGTGCATCCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGA 754
           | || || ||   |  ||||||    || |||||   |||||| ||  |||| |  |  |
Db     414 TTGTCTCCTACTCGCACATCCTGGCGGCCATCCTGAGGATCCAGTCTGGGGAGGGCCGCA 355

Qy     755 GGAAAGCCTTCTGCACCTGCTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAG 814
           | || ||||||| |||||||| |||||||||||| ||| | |||||||| | |||||  |
Db     354 GAAAGGCCTTCTCCACCTGCTCCTCCCACCTCTGCGTGGTGGGACTCTTCTTTGGCAGTG 295

Qy     815 CCATTATCATGTATGTTGGACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCC 874
           |||| ||||||||  | |  |||| |   |  | ||  ||||||| | ||    |
Db     294 CCATCATCATGTACATGGCCCCCAAGTCCCGCCATCCTGAGGAGCAGCAAAAGGTCTTTT 235

Qy     875 TGCTGTTTCACAGCCTCTTTAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACT 934
           | || ||| |||| | || || || |  || || ||||| || |  || || ||||||
Db     234 TTCTATTTTACAGTTTTTTCAACCCAACACTTAACCCCCTGATTTACAGCCTGAGGAACG 175

Qy     935 CAGAAGTGAAGAATACTTTGAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGA 987
           ||| || |||  | |  ||| |||||  |||||    |||||   ||  | |
Db     174 GAGAGGTCAAGGGTGCCCTGAGGAGAGCACTGGGCAAGGAAAGTCATTCCTAA 122


RESULT 15
AAS08660
ID   AAS08660 standard; cDNA; 2282 BP.
XX
AC   AAS08660;
XX
DT   17-DEC-2001  (first entry)
XX
DE   Human cDNA encoding G-protein coupled receptor like protein, GPCR #13.
XX
KW   Human; G-protein coupled receptor like protein; GPCR; immunogen;
KW   ophthalmic disease; neurological disease; Alzheimer's disease;
KW   Parkinson's disease; immunological disorder; HIV; candidiasis;
KW   human immunodeficiency virus; autoimmune disorder; multiple sclerosis;
KW   systemic lupus erythematosus; rheumatoid arthritis; platelet disorder;
KW   thrombocytopaenia; aplastic anaemia; inflammatory disorder;
KW   septic shock; systemic inflammatory response syndrome; SIRS;
KW   hormonal dysfunction; cancer; atherosclerosis; wound;
KW   tissue regeneration; haemophilia; leukaemia; reperfusion injury;
KW   psoriasis; diabetes; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1257..2189
FT                   /*tag= a
FT                   /product= "GPCR #13"
XX
PN   WO200153454-A2.
XX
PD   26-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-US34983.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   20-JUN-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   31-AUG-2000; 2000US-0653450.
PR   04-DEC-2000; 2000US-0729739.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Yamazaki V,  Tang YT,  Liu C,  Zhou P,  Wang D,  Zhang J,  Ren F;
PI   Asundi V,  Drmanac RT;
XX
DR   WPI; 2001-442255/47.
DR   P-PSDB; AAU04582.
XX
PT   New G-protein-coupled receptor-like polypeptides and polynucleotides,
PT   useful for treating diseases of ophthalmic, neurological, immunological
PT   and nephritic systems and hormonal dysfunction, cancer, atherosclerosis
PT   and diabetes -
XX
PS   Claim 1; Page 256-258; 259pp; English.
```

XX
CC The sequence encodes a human G-protein coupled receptor (GPCR)-like
CC protein. The GPCR-like polypeptides and polynucleotides are useful for
CC the treatment of diseases of ophthalmic, neurological (e.g Alzheimer's
CC disease and Parkinson's disease, immunological (e.g HIV infection and
CC candidiasis), autoimmune disorders (e.g multiple sclerosis, systemic
CC lupus erythematosus and rheumatoid arthritis), platelet disorders (e.g
CC thrombocytopaenia and aplastic anaemia), inflammatory disorders (e.g.
CC septic shock and systemic inflammatory response syndrome, SIRS) and
CC nephritic systems. They may also be used to treat hormonal dysfunction,
CC cancer, atherosclerosis, wounds, tissue regeneration, haemophilia,
CC leukaemias, reperfusion injury, psoriasis and diabetes. Numerous examples
CC of each type of disorder are given in the specification. Anti-GPCR-like
CC protein antibodies are useful for detecting or quantitating the
CC polypeptide in tissue. The polypeptides can also be used as molecular
CC weight markers and as a food supplement.
XX
SQ Sequence 2282 BP; 583 A; 573 C; 428 G; 698 T; 0 other;

Query Match 53.1%; Score 538.2; DB 22; Length 2282;
Best Local Similarity 73.6%; Pred. No. 7.2e-151;
Matches 712; Conservative 0; Mismatches 253; Indels 2; Gaps 2;

```
Qy      31 CTTGATTTCCTCTGTCATAGGGATATGGGGGACAATATAACATCCATCACAGAGTTCCTC 90
           ||| ||| | | |  ||| || |||| | | |||   |||    ||||||||||||||
Db    1233 CTTTGTTTGTTTTTTAGTAGTGAAATGGTGAAAAATCAGACAATGGTCACAGAGTTCCTC 1292

Qy      91 CTACTGGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 150
           ||||||||||||| | | ||||||||||||||||||||||||||||||||||||||||||
Db    1293 CTACTGGGATTTCTCCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 1352

Qy     151 TTCTACGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCC 210
           ||||| ||||||||||||||||||||| |||||||| ||||||||||||||||||||||||
Db    1353 TTCTATGTCTTCACCCTGCTGGGGAATGGGACCATCCTGGGGCTCATCTCACTGGACTCC 1412

Qy     211 AGACTGCACGCCCCC-TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTAC 269
           ||||| ||| ||||| ||||||||||||||||||||||||| |||||| ||||||||||
Db    1413 AGACTCCACACCCCCATGTACTTCTTCCTCTCACACCTGGCCGTCGTCAACATCGCCTAT 1472

Qy     270 GCCTGCAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCC 329
           |||||||||||| ||||||| ||||||||||||||||||||||||||||||||||||||||
Db    1473 GCCTGCAACACAGTGCCCCAGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCC 1532

Qy     330 TTTGCGGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTC 389
           ||||| ||| |||||| ||||||||||| |||| | ||||||    || ||||| |||
Db    1533 TTTGCTGGCTGCATGACACAGACCTTTCTCTTTTTGAGTTTTGCACATACTGAATGCCTC 1592

Qy     390 CTCCTGGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTG 449
           || ||||| |||||||||| |||| ||||||||||||||||||| ||||||||||
Db    1593 CTGTTGGTGCTGATGTCCTACGATCGGTACGTGGCCATCTGCCACCCTCTCCGATATTTC 1652

Qy     450 GCCATCATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTC 509
             |||||||||||||| |||||||||||| || ||  | ||||||||||   |||   |
Db    1653 ATCATCATGACCTGGAAAGTCTGCATCACTCTGGCCATCACTTCCTGGACATGTGGCTCC 1712

Qy     510 CTTTTATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAA 569
           || |  |  || | ||| |     |  | |||    | ||||| ||| |||| |   ||
Db    1713 CTCCTGGCTATGGTCCATGTGAGCCTCATCCTAAGACTGCCCTTTTGTGGGCCTCGTGAA 1772

Qy     570 ATTTATCACTTTTTTTTGTGAAATCTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCA 629
           ||  | ||| || || ||||||||| || |||| |||| || ||||||||| ||||||
Db    1773 ATCAACCAC-TTCTTCTGTGAAATCCTGTCTGTCCTCAGGCTGGCCTGTGCTGATACCTG 1831

Qy     630 CATCAATGAGAACATGGTCTTGGCCGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCAC 689
             ||||  ||    |  |||| || |     || |     ||||||||||  | | |
Db    1832 GCTCAACCAGGTGGTCATCTTTGCAGCCTGCATGTTCATCCTGGTGGGACCACTCTGCCT 1891

Qy     690 AATTGTAGTTTCATATATGTGCATCCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGT 749
             |  | || || ||      ||||||    || |||||   |||||| ||  |||| |
Db    1892 GGTGCTGGTCTCCTACTCACACATCCTGGCGGCCATCCTGAGGATCCAGTCTGGGGAGGG 1951

Qy     750 TCAGAGGAAAGCCTTCTGCACCTGCTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGG 809
            |  || .|| ||||||| |||||||| |||||||||||| ||  | |||||||| | |||

Db    1952 CCGCAGAAAGGCCTTCTCCACCTGCTCCTCCCACCTCTGCGTAGTGGGACTCTTCTTTGG 2011

Qy     810 CACAGCCATTATCATGTATGTTGGACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATA 869
           ||  |||||  |||||||  | |  || |  |    |  | ||  |||||||| ||||
Db    2012 CAGCGCCATCGTCATGTACATGGCCCCTAAGTCCCGCCATCCTGAGGAGCAGCAGAAGGT 2071

Qy     870 TCTCCTGCTGTTTCACAGCCTCTTTAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAG 929
             ||  | || ||| ||||    || || || ||||| || ||||| || |  |  || ||
Db    2072 CCTTTTTCTATTTTACAGTTCTTTCAACCCGATGCTAAACCCCCTGATTTACAACCTGAG 2131

Qy     930 GAACTCAGAAGTGAAGAATACTTTGAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGAAA 989
           |||   ||| || |||  | |  ||| ||||| ||| |    ||||||   || | | |
Db    2132 GAATGTAGAGGTCAAGGGTGCCCTGAGGAGAGCACTGTGCAAGGAAAGTCATTCCTAAGA 2191

Qy     990 AGGATTA 996
             | | |
Db    2192 GGTGTGA 2198
```

Search completed: May 2, 2003, 05:01:21
Job time : 293 secs

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 04:52:26 ; Search time 1543 Seconds
(without alignments)
10643.048 Million cell updates/sec

Title: US-09-898-586-23
Perfect score: 1014
Sequence: 1 taaacacttctcctaaacca.....tatggcattgtgactgacaa 1014

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, , and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	532.8	52.5	2021	11	BC016940	BC016940 Homo sapi
c	2	432	42.6	456	9	AI339565	AI339565 qk67e07.x
c	3	429.4	42.3	443	13	BG984967	BG984967 IL5-CN006
	4	417.4	41.2	456	14	BQ301355	BQ301355 QV1-HB003
c	5	409	40.3	433	9	AI339573	AI339573 qk67f03.x
	6	398.4	39.3	619	9	AI148854	AI148854 qc70a02.x
	7	394.8	38.9	705	11	AF327904	AF327904 Homo sapi
	8	394.8	38.9	751	12	BE856570	BE856570 7f64e08.x
	9	380.6	37.5	738	9	AI079550	AI079550 oz04f08.x
	10	370.4	36.5	732	12	BF116115	BF116115 7n77g04.x
	11	359	35.4	580	10	BE543056	BE543056 601068985
c	12	355	35.0	367	9	AI340119	AI340119 qk64f08.x
	13	345.6	34.1	426	14	H62441	H62441 yr79d08.r1
	14	345.6	34.1	623	13	BM670360	BM670360 UI-E-DW1-
c	15	337.4	33.3	361	9	AI285267	AI285267 ql38c08.x
	16	332.6	32.8	690	10	BE179276	BE179276 RC1-HT061
	17	331	32.6	439	14	H62445	H62445 yr79e08.r1
	18	320	31.6	441	13	BI020419	BI020419 CM3-MT029
	19	305.2	30.1	499	10	AW299289	AW299289 xs38f05.x
c	20	293.8	29.0	332	10	BE182734	BE182734 RC3-HT065
	21	272.4	26.9	788	17	BH111304	BH111304 RPC1-24-3
c	22	267.4	26.4	432	10	BE174029	BE174029 QV1-HT057
c	23	265.2	26.2	671	12	BF974217	BF974217 602243851
	24	242.8	23.9	520	17	AQ600752	AQ600752 HS_5346_B
	25	232	22.9	376	12	BE926655	BE926655 QV1-BT063
c	26	222.6	22.0	449	14	H39851	H39851 yp01h11.r1
	27	214.4	21.1	432	9	AA962054	AA962054 oq79g01.s
	28	211.8	20.9	760	12	BG169291	BG169291 602321020
	29	206.8	20.4	1394	11	AK017036	AK017036 Mus muscu
	30	206.8	20.4	3063	11	AK016560	AK016560 Mus muscu
	31	203	20.0	445	14	H39853	H39853 yp01h12.s1
	32	183.6	18.1	226	9	AI084964	AI084964 ow86f03.s
c	33	180	17.8	231	13	BG984943	BG984943 IL5-CN006
	34	174.4	17.2	796	12	BG197640	BG197640 RST17016
	35	174	17.2	762	12	BG193339	BG193339 RST12467
	36	173.6	17.1	495	12	BF841098	BF841098 PM2-HT035
	37	173.4	17.1	661	13	BG924888	BG924888 HNC36-1-B
c	38	162.8	16.1	445	14	BM712072	BM712072 UI-E-DW1-
c	39	162.4	16.0	428	14	BM707299	BM707299 UI-E-CR1-
	40	160.8	15.9	642	17	AZ969227	AZ969227 2M0241J24
	41	160.4	15.8	1501	11	AK016338	AK016338 Mus muscu
	42	160.2	15.8	732	12	BG068751	BG068751 H3068H10-
	43	156.4	15.4	786	12	BG204873	BG204873 RST24289
	44	155.6	15.3	915	14	BQ888002	BQ888002 AGENCOURT
	45	153.8	15.2	632	17	AZ765752	AZ765752 1M0562E19

ALIGNMENTS

RESULT 1
BC016940
LOCUS BC016940 2021 bp mRNA linear HTC 09-NOV-2001
DEFINITION Homo sapiens, Similar to olfactory receptor, family 2, subfamily A, member 4, clone IMAGE:4424116, mRNA.
ACCESSION BC016940
VERSION BC016940.1 GI:16877381
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2021)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 27 Row: 1 Column: 22
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, GenomeScan gene prediction, Similarity but not identity to protein
This clone has the following problem: frame shifted.

```
FEATURES             Location/Qualifiers
     source          1. .2021
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4424116"
                     /tissue_type="Kidney, hypernephroma"
                     /clone_lib="NIH_MGC_89"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
BASE COUNT      470 a    546 c    427 g    578 t
ORIGIN
```

Query Match 52.5%; Score 532.8; DB 11; Length 2021;
Best Local Similarity 73.5%; Pred. No. 1.9e-134;
Matches 706; Conservative 0; Mismatches 252; Indels 2; Gaps 2;

```
Qy      38 TCCTCTGTCATAGGGATATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACTGG 97
           || |    |  ||| || |||| | | |||   |||    ||||||||||||||||||||
Db     412 TCATAACTATTAGTGAAATGGTGAAAAATCAGACAATGGTCACAGAGTTCCTCCTACTGG 471

Qy      98 GATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACG 157
           |||||| | | ||||||||||||||||||||||||||||||||||||||||||||||| |
Db     472 GATTTCTCCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTATG 531

Qy     158 TCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGC 217
           |||||||||||||||||| |||||||| |||||||||||||||||||||||||||||| |
Db     532 TCTTCACCCTGCTGGGGAATGGGACCATCCTGGGGCTCATCTCACTGGACTCCAGACTCC 591

Qy     218 ACGCCCCC-TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCA 276
           || ||||| |||||||||||||||||||||||| |||||| |||||||||| |||||||
Db     592 ACACCCCCATGTACTTCTTCCTCTCACACCTGGCCGTCGTCAACATCGCCTATGCCTGCA 651

Qy     277 ACACGGTGCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGG 336
           |||| ||||||| ||||||||||||||||||||||||||||||||||||||||||||| |
Db     652 ACACAGTGCCCCAGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCTG 711

Qy     337 GCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGG 396
           || ||||||   |||||||||| |||| | ||||||    || ||||| |||||  |||
Db     712 GCTGCATGACATAGACCTTTCTCTTTTTGAGTTTTGCACATACTGAATGCCTCCTGTTGG 771

Qy     397 TGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCA 456
           || |||||||||| |||| ||||||||||||||||||||| ||||||||||   |||||
Db     772 TGCTGATGTCCTACGATCGGTACGTGGCCATCTGCCACCCTCTCCGATATTTCATCATCA 831

Qy     457 TGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTAT 516
           ||||||||| |||||||||||| || ||  | ||||||||||   |||  |||  |
Db     832 TGACCTGGAAAGTCTGCATCACTCTGGCCATCACTTCCTGGACATGTGGCTCCCTCCTGG 891

Qy     517 CCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATC 576
           |  || | ||| |     |  | |||    | ||||| ||| |||| |   ||||  | |
Db     892 CTATGGTCCATGTGAGCCTCATCCTAAGACTGCCCTTTTGTGGGCCTCGTGAAATCAACC 951

Qy     577 ACTTTTTTTTGTGAAATCTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAAT 636
           || || || |||||||| || |||| |||  || |||||||| ||||||    ||||
Db     952 AC-TTCTTCTGTGAAATCCTGTCTGTCCTCAGGCTGGCCTGTGCTGATACCTGGCTCAAC 1010

Qy     637 GAGAACATGGTCTTGGCCGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTA 696
            ||    |  |||| || |     || |     ||||||||||  | | |    | |
Db    1011 CAGGTGGTCATCTTTGCAGCCTGCATGTTCATCCTGGTGGGACCACTCTGCCTGGTGCTG 1070

Qy     697 GTTTCATATATGTGCATCCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGG 756
           || || ||      ||||||    || |||||   |||||| ||  |||| |  |  ||
```

```
Db    1071 GTCTCCTACTCACACATCCTGGCGGCCATCCTGAGGATCCAGTCTGGGGAGGGCCGCAGA 1130

Qy     757 AAAGCCTTCTGCACCTGCTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCC 816
           || ||||||| |||||||| |||||||||||| ||  | |||||||| | |||||  |||
Db    1131 AAGGCCTTCTCCACCTGCTCCTCCCACCTCTGCGTAGTGGGACTCTTCTTTGGCAGCGCC 1190

Qy     817 ATTATCATGTATGTTGGACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTG 876
           ||  |||||||  | |  || |   |   | ||  |||||||| ||||    ||  |
Db    1191 ATCGTCATGTACATGGCCCCTAAGTCCCGCCATCCTGAGGAGCAGCAGAAGGTCCTTTTT 1250

Qy     877 CTGTTTCACAGCCTCTTTAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCA 936
           || ||| ||||    || || || ||||| || ||||| || |  |  || |||||   |
Db    1251 CTATTTTACAGTTCTTTCAACCCGATGCTAAACCCCCTGATTTACAACCTGAGGAATGTA 1310

Qy     937 GAAGTGAAGAATACTTTGAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTA 996
           || || |||  | |  ||| ||||| ||| |    |||||   ||  | | | |  | |
Db    1311 GAGGTCAAGGGTGCCCTGAGGAGAGCACTGTGCAAGGAAAGTCATTCCTAAGAGGTGTGA 1370
```

RESULT 2
AI339565/c

```
LOCUS       AI339565                456 bp    mRNA    linear   EST 13-FEB-1999
DEFINITION  qk67e07.x1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1874052 3'
            similar to SW:OLF7_MOUSE P34984 OLFACTORY RECEPTOR-LIKE PROTEIN K7
            ;, mRNA sequence.
ACCESSION   AI339565
VERSION     AI339565.1  GI:4076492
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 456)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck,.M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 669   Std Error: 0.00
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .456
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1874052"
                     /clone_lib="NCI_CGAP_Co8"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; 1st strand cDNA was prepared from
                     colon adenocarcinoma, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector.  Library is normalized. Library was constructed by
                     Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      136 a     93 c    149 g     78 t
ORIGIN
```

Query Match 42.6%; Score 432; DB 9; Length 456;
Best Local Similarity 99.8%; Pred. No. 3.7e-107;
Matches 443; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy     143 TCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCAC 202
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    456 TCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCAC 397

Qy    203 TGGACTCCAGACTGCACGCCCCC-TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACA 261
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    396 TGGACTCCAGACTGCACGCCCCCATGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACA 337

Qy    262 TCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGC 321
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    336 TCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGC 277

Qy    322 CCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAG 381
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    276 CCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAG 217

Qy    382 AATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCC 441
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    216 AATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCC 157

Qy    442 GATATTTGGCCATCATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCA 501
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    156 GATATTTGGCCATCATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCA 97

Qy    502 CTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGC 561
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     96 CTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGC 37

Qy    562 CCCAGAAAATTTATCACTTTTTTT 585
          |||||||||||||||||||||||||
Db     36 CCCAGAAAATTTATCACTTTTTTT 13


RESULT 3
BG984967/c
LOCUS       BG984967                443 bp    mRNA    linear   EST 12-JUN-2001
DEFINITION  IL5-CN0068-100401-434-d09 CN0068 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG984967
VERSION     BG984967.1  GI:14387702
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL5&t2=IL5-CN0068-
            100401-434-d09&t3=2001-04-10&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 363.
FEATURES             Location/Qualifiers
     source          1. .443
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="CN0068"
                     /dev_stage="Adult"
                     /note="Organ: colon_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT      136 a     94 c    131 g     81 t      1 others
ORIGIN

  Query Match             42.3%;  Score 429.4;  DB 13;  Length 443;
  Best Local Similarity   99.3%;  Pred. No. 1.9e-106;
  Matches  441;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy    261 ATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAG 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    443 ATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAG 384

Qy    321 CCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACA 380
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    383 CCCATCTCCTNGCGGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACA 324

Qy    381 GAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTC 440
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    323 GAATGTCTCCTCCTGGTGGTGCTGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTC 264

Qy    441 CGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACC 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263 CGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACC 204

Qy    501 ACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGG 560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 ACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGG 144

Qy    561 CCCCAGAAAATTTATCACTTTTTTTTGTGAAATCTTGGCTGTTCTCAAACTTGCCTGTGC 620
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    143 CCCCAGAAAATTTATCAC-TTTTTTTGTGAAATCTTGGCTGTTCTCAAACTTGCCTGTGC 85

Qy    621 AGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGCAATTTCTGGGCTGGTGGGACC 680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     84 AGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGCAATTTCTGGGCTGGTGGGACC 25

Qy    681 CTTGTCCACAATTGTAGTTTCATA 704
          ||||||||||||||||||||||||
Db     24 CTTGTCCACAATTGTAGTTTCATA 1


RESULT 4
BQ301355
LOCUS       BQ301355                456 bp    mRNA    linear   EST 16-MAY-2002
DEFINITION  QV1-HB0037-020201-589-d09 HB0037 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BQ301355
VERSION     BQ301355.1  GI:20816877
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 456)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
```

```
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV1&t2=QV1-HB0037-
            020201-589-d09&t3-2001-02-02&t4=1)
            Seq primer: puc 18 forward.
FEATURES              Location/Qualifiers
     source           1. .456
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone_lib="HB0037"
                      /dev_stage="Adult"
                      /note="Organ: bocio_tumor; Vector: puc18; Site_1: SmaI;
                      Site_2: SmaI; A mini-library was made by cloning products
                      derived from  ORESTES PCR (U.S. Letters Patent application
                      No. 196,716 -  Ludwig Institute for Cancer Research)
                      profiles into  the pUC 18 vector. Reverse transcription of
                      tissue mRNA and cDNA amplification were performed under
                      low stringency conditions."
BASE COUNT       97 a    120 c    101 g    138 t
ORIGIN

  Query Match             41.2%;  Score 417.4;  DB 14;  Length 456;
  Best Local Similarity   98.2%;  Pred. No. 3.6e-103;
  Matches  443;  Conservative    0;  Mismatches    6;  Indels    2;  Gaps    2;

Qy     417 TACGTGGCCATCTGCCACCCCCTCCGA-TATTTGGCCATCATGACCTGGAGAGTCTGCAT 475
           ||||  | ||||||||||||||||||| |||||||||||||||||||||||||||||||
Db       4 TACGGTGGCATCTGCCACCCCCTCCGAGTATTTGGCCATCATGACCTGGAGAGTCTGCAT 63

Qy     476 CACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTT 535
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64 CACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTT 123

Qy     536 ACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAAATCT 595
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     124 ACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAATCT 182

Qy     596 TGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCG 655
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 TGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCG 242

Qy     656 GAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCC 715
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 GAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCC 302

Qy     716 TCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCT 775
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 TCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCT 362

Qy     776 TCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGAC 835
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 TCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGAC 422

Qy     836 CCAGATATGGGAACCCCAAGGAGCAGAAGAA 866
           |||||||||||||||| ||||| ||||| |
Db     423 CCAGATATGGGAACCCCGAGGAGGAGAAGCA 453


RESULT 5
AI339573/c
LOCUS       AI339573                433 bp    mRNA    linear   EST 13-FEB-1999
DEFINITION  qk67f03.x1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1874045 3'
            similar to SW:OLF7_MOUSE P34984 OLFACTORY RECEPTOR-LIKE PROTEIN K7
            ;, mRNA sequence.
ACCESSION   AI339573
VERSION     AI339573.1  GI:4076500
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 433)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 677    Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 411.
FEATURES              Location/Qualifiers
     source           1. .433
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:1874045"
                      /clone_lib="NCI_CGAP_Co8"
                      /tissue_type="adenocarcinoma"
                      /lab_host="DH10B"
                      /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                      modified polylinker; 1st strand cDNA was prepared from
                      colon adenocarcinoma, and was then primed with a Not I -
                      oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                      RI adaptors (Pharmacia), digested with Not I and cloned
                      into the Not I and Eco RI sites of the modified pT7T3
                      vector.  Library is normalized. Library was constructed by
                      Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      127 a     91 c    139 g     76 t
ORIGIN

  Query Match             40.3%;  Score 409;  DB 9;  Length 433;
  Best Local Similarity   99.8%;  Pred. No. 7.1e-101;
  Matches  420;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy     166 CTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC 225
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     433 CTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC 374

Qy     226 -TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTG 284
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     373 ATGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTG 314

Qy     285 CCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATG 344
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     313 CCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATG 254

Qy     345 ATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATG 404
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     253 ATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATG 194

Qy     405 TCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGG 464
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     193 TCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGG 134

Qy     465 AGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATT 524
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     133 AGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATT 74

Qy     525 CATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTT 584
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      73 CATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTT 14

Qy     585 T 585
```

```
Db      13 T 13


RESULT 6
AI148854
LOCUS       AI148854                619 bp    mRNA    linear   EST 28-OCT-1998
DEFINITION  qc70a02.x1 Soares_placenta_8to9weeks_2NbHP8to9W Homo sapiens cDNA
            clone IMAGE:1714922 3' similar to SW:OLF7_MOUSE P34984 OLFACTORY
            RECEPTOR-LIKE PROTEIN K7 ;, mRNA sequence.
ACCESSION   AI148854
VERSION     AI148854.1  GI:3677323
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 619)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1098   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 432.
FEATURES             Location/Qualifiers
     source          1. .619
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1714922"
                     /clone_lib="Soares_placenta_8to9weeks_2NbHP8to9W"
                     /dev_stage="two placentae: one from 8 weeks and another
                     from 9 weeks post conception"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGATTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo."
BASE COUNT      112 a    191 c    131 g    183 t      2 others
ORIGIN

  Query Match             39.3%;  Score 398.4;  DB 9;  Length 619;
  Best Local Similarity   80.3%;  Pred. No. 6.5e-98;
  Matches  490;  Conservative    0;  Mismatches  118;  Indels    2;  Gaps    2;

Qy      31 CTTGATTTCCTCTGTCATAGGGATATGGGGGACAATATAACATCCATCACAGAGTTCCTC 90
           ||| ||| || | |  ||| || |||| | | |||   |||    ||||||||||||||
Db       2 CTTTGTTTGCTTTTTAGTAGTGAAATGGTGAAAAATCAGACAATGGTCACAGAGTTCCTC 61

Qy      91 CTACTGGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 150
           |||||||||||||| | | ||||||||||||||||||||||||||||||||||||||||
Db      62 CTACTGGGATTTCTCCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 121

Qy     151 TTCTACGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCC 210
           ||||| ||||||||||||||||||||| |||||||| ||||||||||||||||||||||||
Db     122 TTCTATGTCTTCACCCTGCTGGGGAATGGGACCATCCTGGGGCTCATCTCACTGGACTCC 181

Qy     211 AGACTGCACGCCCCC-TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTAC 269
           ||||| ||| ||||| |||||||||||||||||||||||||| |||||| || |||||||
Db     182 AGACTCCACACCCCCATGTACTTCTTCCTCTCACACCTGGCCGTCGTCAACGTCGCCTAT 241

Qy     270 GCCTGCAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCC 329
           |||||||||||| ||||||| ||||||||||||||||||||||||||||||||||||||||
Db     242 GCCTGCAACACAGTGCCCCAGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCC 301

Qy     330 TTTGCGGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTC 389
           ||||| ||| |||||||   |||||||||| |||| | ||||||    || ||||| |||
Db     302 TTTGCTGGCTGCATGACATAGACCTTTCTCTTTTTGAGTTTTGCACATACTGAATGCCTC 361

Qy     390 CTCCTGGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTG 449
           || ||||| |||||||||| |||| |||||||||||||||||||||| ||||||||||
Db     362 CTGTTGGTGCTGATGTCCTACGATCGGTACGTGGCCATCTGCCACCCTCTCCGATATTTC 421

Qy     450 GCCATCATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTC 509
             |||||||||||| | |||||||||||| || ||  | ||||||||||   |||   |
Db     422 ATCATCATGACCTGNAAAGTCTGCATCACTCTGGCCATCACTTCCTGGACATGTGGCTCC 481

Qy     510 CTTTTATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAA 569
           || |  |  ||   ||| |     |  | |||    | ||||| ||| |||| |   ||
Db     482 CTCCTGGCTATGGNCCATGTGAGCCTCATCCTAAGACTGCCCTTTTGTGGGCCTCGTGAA 541

Qy     570 ATTTATCACTTTTTTTTGTGAAATCTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCA 629
           ||  | ||| || || |||||||||| || |||| ||||  || | ||||||  ||||||
Db     542 ATCAACCAC-TTCTTCTGTGAAATCCTGTCTGTCCTCAGGCTGGGCTGTGCTGATACCTG 600

Qy     630 CATCAATGAG 639
             ||||  ||
Db     601 GCTCAACCAG 610


RESULT 7
AF327904
LOCUS       AF327904                705 bp    mRNA    linear   HTC 13-JUL-2001
DEFINITION  Homo sapiens FKSG35 (FKSG35) mRNA, complete cds.
ACCESSION   AF327904
VERSION     AF327904.1  GI:12276181
KEYWORDS    HTC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 705)
  AUTHORS   Wang,Y.-G. and Gong,L.
  TITLE     Cloning of FKSG35, a novel gene located on human chromosome 7
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 705)
  AUTHORS   Wang,Y.-G.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-DEC-2000) Beijing Fengkesheng Function Gene
            Technology Ltd., 4 Tou Tiao Lu Chang Street, Xuanwu District,
            Beijing 100050, P.R. China
FEATURES             Location/Qualifiers
     source          1. .705
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q33-q35"
                     /tissue_type="liver"
     gene            1. .705
                     /gene="FKSG35"
     CDS             133. .429
                     /gene="FKSG35"
                     /codon_start=1
                     /product="FKSG35"
                     /protein_id="AAG50283.1"
                     /db_xref="GI:12276182"
                     /translation="MVKNQTMVTEFLLLGFLLGPRIQMLLFGLFSLFYVFTLLGNGTI
                     LGLISLDSRLHTPMYFFLSHLAVVNIAYACNTVPQMLVNLLHPAKPISFAGCMT"
BASE COUNT      156 a    201 c    132 g    216 t
ORIGIN

  Query Match             38.9%;  Score 394.8;  DB 11;  Length 705;
  Best Local Similarity   80.9%;  Pred. No. 6.7e-97;
  Matches  484;  Conservative    0;  Mismatches  112;  Indels    2;  Gaps    2;

Qy      31 CTTGATTTCCTCTGTCATAGGGATATGGGGGACAATATAACATCCATCACAGAGTTCCTC 90
           ||| ||| | | |  ||| || |||| | | |||   |||    ||||||||||||||
```

```
Db    109 CTTTGTTTGTTTTTTAGTAGTGAAATGGTGAAAAATCAGACAATGGTCACAGAGTTCCTC 168

Qy     91 CTACTGGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 150
          |||||||||||||| | | ||||||||||||||||||||||||||||||||||||||||||
Db    169 CTACTGGGATTTCTCCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 228

Qy    151 TTCTACGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCC 210
          ||||| |||||||||||||||||||| |||||||| ||||||||||||||||||||||||
Db    229 TTCTATGTCTTCACCCTGCTGGGGAATGGGACCATCCTGGGGCTCATCTCACTGGACTCC 288

Qy    211 AGACTGCACGCCCCC-TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTAC 269
          ||||| ||| ||||| ||||||||||||||||||||||||| |||||| ||||||||||
Db    289 AGACTCCACACCCCCATGTACTTCTTCCTCTCACACCTGGCCGTCGTCAACATCGCCTAT 348

Qy    270 GCCTGCAACACGGTGCCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCC 329
          |||||||||||| ||||||| ||||||||||||||||||||||||||||||||||||||||
Db    349 GCCTGCAACACAGTGCCCCAGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCC 408

Qy    330 TTTGCGGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTC 389
          ||||| ||| |||||||   |||||||||| |||| | ||||||    || ||||| |||
Db    409 TTTGCTGGCTGCATGACATAGACCTTTCTCTTTTTGAGTTTTGCACATACTGAATGCCTC 468

Qy    390 CTCCTGGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTG 449
          ||  ||||| |||||||||| |||| ||||||||||||||||||||| ||||||||||
Db    469 CTGTTGGTGCTGATGTCCTACGATCGGTACGTGGCCATCTGCCACCCTCTCCGATATTTC 528

Qy    450 GCCATCATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTC 509
          ||||||||||||||| |||||||||||| || ||  | ||||||||||   |||  |
Db    529 ATCATCATGACCTGGAAAGTCTGCATCACTCTTGCCATCACTTCCTGGACATGTGGCTCC 588

Qy    510 CTTTTATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAA 569
          ||  |  |  || | ||| |     |  | |||    | ||||| ||| |||| |  ||
Db    589 CTCCTGGCTATGGTCCATGTGAGCCTCATCCTAAGACTGCCCTTTTGTGGGCCTCGTGAA 648

Qy    570 ATTTATCACTTTTTTTTTGTGAAATCTTGGCTGTTCTCAAACTTGCCTGTGCAGATACC 627
          ||  | ||| || || ||||||| | || |||| |   | || |||||||| ||||||
Db    649 ATCAACCAC-TTCTTCTGTGAAAACCTGTCTGTCCCTCAGCTGGCCTGTGCTGATACC 705
```

RESULT 8
BE856570

LOCUS	BE856570 751 bp mRNA linear EST 29-SEP-2000
DEFINITION	7f64e08.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone IMAGE:3299462 3' similar to TR:O95047 O95047 WUGSC:H_DJ0988G15.2 PROTEIN. ;, mRNA sequence.
ACCESSION	BE856570
VERSION	BE856570.1 *GI:10369724*
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 751)
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	*Unpublished* (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: -40UP from Gibco High quality sequence stop: 449.

FEATURES	*Location/Qualifiers*
source	1. .751 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:3299462" /clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1" /lab_host="DH10B" /note="Organ: pooled; Vector: *pT7T3D-Pac* (*Pharmacia*) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from five normalized libraries were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and cloneIDs: Soares *NbHSF pool 1*: 309384-310919, 323208-325895 Soares Nb2HP pool 1: 145032-147335, 147720-148103, 148872-149255, 15002 - 150407, 151176-152327 Soares Nb2HF8-9W pool 1: 758280-760583, 772104-774407 Soares NbHPA pool 1: 304776-306311, 320136-322823, 326280-326663 Soares NbHOT pool 1: 723720-726407, 739080-740999 Subtraction by Bento Soares and M. Fatima Bonaldo."

```
BASE COUNT      170 a    206 c    137 g    238 t
ORIGIN

  Query Match           38.9%;  Score 394.8;  DB 12;  Length 751;
  Best Local Similarity  80.9%;  Pred. No. 6.9e-97;
  Matches  484;  Conservative    0;  Mismatches  112;  Indels    2;  Gaps    2;

Qy     31 CTTGATTTCCTCTGTCATAGGGATATGGGGGACAATATAACATCCATCACAGAGTTCCTC 90
          |||  |||  | | |  ||| || |||| | | |||   |||    ||||||||||||||
Db    155 CTTTGTTTGTTTTTTAGTAGTGAAATGGTGAAAAATCAGACAATGGTCACAGAGTTCCTC 214

Qy     91 CTACTGGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 150
          |||||||||||| | | ||||||||||||||||||||||||||||||||||||||||||
Db    215 CTACTGGGATTTCTCCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 274

Qy    151 TTCTACGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCC 210
          ||||| |||||||||||||||||| |||||||| ||||||||||||||||||||||||
Db    275 TTCTATGTCTTCACCCTGCTGGGGAATGGGACCATCCTGGGGCTCATCTCACTGGACTCC 334

Qy    211 AGACTGCACGCCCCC-TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTAC 269
          ||||| ||| ||||| ||||||||||||||||||||||||| |||||| ||||||||||
Db    335 AGACTCCACACCCCCATGTACTTCTTCCTCTCACACCTGGCCGTCGTCAACATCGCCTAT 394

Qy    270 GCCTGCAACACGGTGCCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCC 329
          ||||||||||| ||||||| ||||||||||||||||||||||||||||||||||||||||
Db    395 GCCTGCAACACAGTGCCCCAGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCC 454

Qy    330 TTTGCGGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTC 389
          ||||| ||| ||||||   |||||||||| |||| | ||||||    || ||||| |||
Db    455 TTTGCTGGCTGCATGACATAGACCTTTCTCTTTTTGAGTTTTGCACATACTGAATGCCTC 514

Qy    390 CTCCTGGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTG 449
          ||  ||||| |||||||||| |||| ||||||||||||||||||||| ||||||||||
Db    515 CTGTTGGTGCTGATGTCCTACGATCGGTACGTGGCCATCTGCCACCCTCTCCGATATTTC 574

Qy    450 GCCATCATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTC 509
          |||||||||||||| |||||||||||| || ||  | ||||||||||   |||  |
Db    575 ATCATCATGACCTGGAAAGTCTGCATCACTCTTGCCATCACTTCCTGGACATGTGGCTCC 634

Qy    510 CTTTTATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAA 569
          ||  |  |  || | ||| |     |  | |||    | ||||| ||| |||| |  ||
Db    635 CTCCTGGCTATGGTCCATGTGAGCCTCATCCTAAGACTGCCCTTTTGTGGGCCTCGTGAA 694

Qy    570 ATTTATCACTTTTTTTTTGTGAAATCTTGGCTGTTCTCAAACTTGCCTGTGCAGATACC 627
          ||  | ||| || || ||||||| | || |||| |   | || |||||||| ||||||
Db    695 ATCAACCAC-TTCTTCTGTGAAAACCTGTCTGTCCCTCAGCTGGCCTGTGCTGATACC 751
```

RESULT 9
AI079550

LOCUS	AI079550 738 bp mRNA linear EST 29-SEP-1998
DEFINITION	oz04f08.x1 Soares_fetal_liver_spleen_1NFLS_S1 *Homo sapiens* cDNA clone IMAGE:1674375 3' similar to SW:OLF7_MOUSE P34984 OLFACTORY RECEPTOR-LIKE PROTEIN K7 ;, mRNA sequence.
ACCESSION	*AI079550*
VERSION	AI079550.1 GI:3415801
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 738)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1896   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 447.
FEATURES             Location/Qualifiers
     source          1. .738
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1674375"
                     /clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     This is a subtracted version of the original Soares fetal
                     liver spleen 1NFLS library.   1st strand cDNA was primed
                     with a Pac I - oligo(dT) primer [5'
                     AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      173 a    201 c    135 g    229 t
ORIGIN

  Query Match             37.5%;  Score 380.6;  DB 9;  Length 738;
  Best Local Similarity   80.8%;  Pred. No. 5.2e-93;
  Matches  468;  Conservative    0;  Mismatches  109;  Indels    2;  Gaps    2;

Qy      31 CTTGATTTCCTCTGTCATAGGGATATGGGGGACAATATAACATCCATCACAGAGTTCCTC 90
           ||| ||| | | |  ||| || |||| | | |||   |||    ||||||||||||||
Db     160 CTTTGTTTGTTTTTTAGTAGTGAAATGGTGAAAAATCAGACAATGGTCACAGAGTTCCTC 219

Qy      91 CTACTGGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 150
           |||||||||||| | | ||||||||||||||||||||||||||||||||||||||||||
Db     220 CTACTGGGATTTCTCCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 279

Qy     151 TTCTACGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCC 210
           ||||| ||||||||||||||||||||| |||||||| ||||||||||||||||||||||||
Db     280 TTCTATGTCTTCACCCTGCTGGGGAATGGGACCATCCTGGGGCTCATCTCACTGGACTCC 339

Qy     211 AGACTGCACGCCCCC-TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTAC 269
           ||||| ||| ||||| |||||||||||||||||||||||||| |||||| ||||||||||
Db     340 AGACTCCACACCCCCATGTACTTCTTCCTCTCACACCTGGCCGTCGTCAACATCGCCTAT 399

Qy     270 GCCTGCAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCC 329
           |||||||||||| ||||||| |||||||||||||||||||||||||||||||||||||||
Db     400 GCCTGCAACACAGTGCCCCAGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCC 459

Qy     330 TTTGCGGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTC 389
           ||||| ||| ||||||   |||||||||| ||||  | ||||||    || ||||| |||
Db     460 TTTGCTGGCTGCATGACATAGACCTTTCTCTTTTTGAGTTTTGCACATACTGAATGCCTC 519

Qy     390 CTCCTGGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTG 449
           ||  ||||| |||||||||| |||| ||||||||||||||||||||| ||||||||||
Db     520 CTGTTGGTGCTGATGTCCTACGATCGGTACGTGGCCATCTGCCACCCTCTCCGATATTTC 579

Qy     450 GCCATCATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTC 509
             |||||||||||||| |||||||||||| || ||  | ||||||||||   |||   |
Db     580 ATCATCATGACCTGGAAAGTCTGCATCACTCTGGCCATCACTTCCTGGACATGTGGCTCC 639

Qy     510 CTTTTATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAA 569
           || |  |  || | ||| |     |  | |||    |  || | |  |||| |   ||
Db     640 CTCCTGGCTATGGTCCATGTGAGCCTCATCCTAAGACTGGCCCTTTTGTGGCCTCGTGAA 699

Qy     570 ATTTATCACTTTTTTTTGTGAAATCTTGGCTGTTCTCAA 608
           ||  | ||| ||| | |||||||| | ||||| |||||
Db     700 ATCAACCAC-TTTCTCTGTGAAATCCTTGCTGTCCTCAA 737


RESULT 10
BF116115
LOCUS       BF116115               732 bp    mRNA    linear   EST 24-OCT-2000
DEFINITION  7n77g04.x1 NCI_CGAP_Ov18 Homo sapiens cDNA clone IMAGE:3570631 3'
            similar to TR:O95047 O95047 WUGSC:H_DJ0988G15.2 PROTEIN. ;, mRNA
            sequence.
ACCESSION   BF116115
VERSION     BF116115.1  GI:10985591
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 732)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento
            Soares, Ph.D. cDNA Library Arrayed by: Christa Prange, The
            I.M.A.G.E. Consortium DNA Sequencing by: Washington University
            Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 472.
FEATURES             Location/Qualifiers
     source          1. .732
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3570631"
                     /clone_lib="NCI_CGAP_Ov18"
                     /tissue_type="fibrotheoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: ovary; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGCGACATTTTTTTTTTTTTTTTT 3'];
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia),  digested with Not I and cloned into the Not
                     I and Eco RI sites  of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo.  "
BASE COUNT      152 a    217 c    145 g    218 t
ORIGIN

  Query Match             36.5%;  Score 370.4;  DB 12;  Length 732;
  Best Local Similarity   83.5%;  Pred. No. 3.2e-90;
  Matches  444;  Conservative    0;  Mismatches   86;  Indels    2;  Gaps    2;

Qy      36 TTTCCTCTGTCATAGGGATATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACT 95
           ||| || | ||| ||||| |||||||| ||| |||    ||||||||||||||||||
Db     200 TTTTCTTTTTCACAGGGAAATGGGGGAAAATCAGACAATGGTCACAGAGTTCCTCCTACT 259

Qy      96 GGGATTTC-CCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCT 154
           |||||||| ||  ||||||||||||||||||||||||||||||||||||||||||||||
Db     260 GGGATTTCTCCTGGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCT 319

Qy     155 ACGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGAC 214
           |  |||||||||||||||||||||| |||| ||||||||||||||||||||||||||||
Db     320 ATATCTTCACCCTGCTGGGGAACGGGGCCATCCTGGGGCTCATCTCACTGGACTCCAGAC 379
```

```
Qy    215 TGCACGCCCCC-TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCT 273
          | ||| ||||| ||||||||||||||||||||||||| ||||||||||||||||||  ||
Db    380 TCCACACCCCCATGTACTTCTTCCTCTCACACCTGGCTGTCGTCGACATCGCCTACACCC 439

Qy    274 GCAACACGGTGCCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTG 333
          |||||||||||||| |||||||| ||||||||||||||||||||||||||||||||||||
Db    440 GCAACACGGTGCCCCAGATGCTGGCGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTG 499

Qy    334 CGGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCC 393
          | ||| |||||| |||||||||||| | ||  | |||||    ||  ||||||||||| |
Db    500 CTGGCTGCATGACGCAGACCTTTCTCTGTTTGAGTTTTGGACACAGCGAATGTCTCCTGC 559

Qy    394 TGGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCA 453
          ||||| |||||||||| |||| ||||||||||||||||||||| |||||||| |  | ||
Db    560 TGGTGCTGATGTCCTACGATCGTTACGTGGCCATCTGCCACCCTCTCCGATACTCCGTCA 619

Qy    454 TCATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTT 513
          ||||||||||||||||||||||||||| || || ||||||  |||   |||   |||
Db    620 TCATGACCTGGAGAGTCTGCATCACCCTGGCCGTCACTTCCCTGACGTGTGGCTCCCTCC 679

Qy    514 TATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCA 565
          |  |  ||   ||| | ||  |  | |||    | ||||||| | |||| ||
Db    680 TGGCTCTGGCCCATGTGGTTCTCATCCTAAGACTGCCCTTCTCTGGGCCTCA 731
```

RESULT 11
BE543056

```
LOCUS       BE543056                580 bp    mRNA    linear   EST 09-AUG-2000
DEFINITION  601068985F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:3455324 5',
            mRNA sequence.
ACCESSION   BE543056
VERSION     BE543056.1  GI:9771701
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 580)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Incyte Genomics, Inc.
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM8441  row: i  column: 21
            High quality sequence start: 3
            High quality sequence stop: 580.
FEATURES             Location/Qualifiers
     source          1. .580
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3455324"
                     /clone_lib="NIH_MGC_12"
                     /tissue_type="cervical carcinoma cell line"
                     /lab_host="DH10B"
                     /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.4 kb. Library prepared by Life
                     Technologies."
BASE COUNT       96 a    189 c    125 g    170 t
ORIGIN
```

```
  Query Match             35.4%;  Score 359;  DB 10;  Length 580;
  Best Local Similarity   80.6%;  Pred. No. 3.8e-87;
  Matches  444;  Conservative    0;  Mismatches  105;  Indels    2;  Gaps    2;

Qy    116 GGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGA 175
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db      1 GGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTATGTCTTCACCCTGCTGGGGA 60

Qy    176 ACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTCT 234
          | ||||||||| ||||||||||||||||||||||||||||| ||| ||||| |||||||||
Db     61 ATGGGACCATCCTGGGGCTCATCTCACTGGACTCCAGACTCCACACCCCCATGTACTTCT 120

Qy    235 TCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGC 294
          |||||||||||||||| |||||| |||||||||| |||||||||| ||||||| |||||
Db    121 TCCTCTCACACCTGGCCGTCGTCAACATCGCCTATGCCTGCAACACAGTGCCCCAGATGC 180

Qy    295 TGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCT 354
          |||||||||||||||||||||||||||||||||||||||||| ||| ||||||   ||||||
Db    181 TGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCTGGCTGCATGACATAGACCT 240

Qy    355 TTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATC 414
          |||| |||| | ||||||    || ||||| |||||  ||||| |||||||||| ||||
Db    241 TTCTCTTTTGAGTTTTGCACATACTGAATGCCTCCTGTTGGTGCTGATGTCCTACGATC 300

Qy    415 TGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCA 474
          |||||||||||||||||||| ||||||||||   |||||||||||||| ||||||||||
Db    301 GGTACGTGGCCATCTGCCACCCTCTCCGATATTTCATCATCATGACCTGGAAAGTCTGCA 360

Qy    475 TCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGT 534
          |||| || ||  | ||||||||||   |||  |||  |  |  || | ||| |
Db    361 TCACTCTGGCCATCACTTCCTGGACATGTGGCTCCCTCCTGGCTATGGTCCATGTGAGCC 420

Qy    535 TACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAAATC 594
          |  | |||    | ||||| ||| |||| |   ||||  | ||| || || |||||||||
Db    421 TCATCCTAAGACTGCCCTTTTGTGGGCCTCGTGAAATCAACCAC-TTCTTCTGTGAAATC 479

Qy    595 TTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCC 654
          || |||| |||| || |||||||| ||||||    ||||  ||    |  |||| ||
Db    480 CTGTCTGTCCTCAGGCTGGCCTGTGCTGATACCTGGCTCAACCAGGTGGTCATCTTTGCG 539

Qy    655 GGAGCAATTTC 665
          |||  |||
Db    540 CCTGCATGTTC 550
```

RESULT 12
AI340119/c

```
LOCUS       AI340119                367 bp    mRNA    linear   EST 29-DEC-1998
DEFINITION  qk64f08.x1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1873767 3'
            similar to SW:OLF7_MOUSE P34984 OLFACTORY RECEPTOR-LIKE PROTEIN K7
            ;, mRNA sequence.
ACCESSION   AI340119
VERSION     AI340119.1  GI:4077046
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 367)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco
            High quality sequence stop: 245.
FEATURES             Location/Qualifiers
     source          1. .367
```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1873767"
/clone_lib="NCI_CGAP_Co8"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from colon adenocarcinoma, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 115 a 73 c 116 g 63 t
ORIGIN

Query Match 35.0%; Score 355; DB 9; Length 367;
Best Local Similarity 100.0%; Pred. No. 3.9e-86;
Matches 355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    231 TTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGG 290
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    367 TTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGG 308

Qy    291 ATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAG 350
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    307 ATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAG 248

Qy    351 ACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTAT 410
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    247 ACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTAT 188

Qy    411 GATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTC 470
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    187 GATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTC 128

Qy    471 TGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTT 530
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    127 TGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTT 68

Qy    531 GTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTT 585
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     67 GTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTT 13
```

RESULT 13
H62441
LOCUS H62441 426 bp mRNA linear EST 06-OCT-1995
DEFINITION yr79d08.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:211503 5' similar to SP:OLF7_MOUSE P34984 OLFACTORY RECEPTOR-LIKE PROTEIN K7 ;, mRNA sequence.
ACCESSION H62441
VERSION H62441.1 GI:1015273
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 426)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1867
High quality sequence stops: 304
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1867 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 304.
FEATURES Location/Qualifiers
source 1. .426
/organism="Homo sapiens"
/db_xref="GDB:3784344"
/db_xref="taxon:9606"
/clone="IMAGE:211503"
/clone_lib="Soares fetal liver spleen 1NFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 75 a 140 c 95 g 111 t 5 others
ORIGIN

Query Match 34.1%; Score 345.6; DB 14; Length 426;
Best Local Similarity 96.0%; Pred. No. 1.6e-83;
Matches 406; Conservative 0; Mismatches 12; Indels 5; Gaps 5;

```
Qy     61 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 60

Qy    121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 120

Qy    181 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTCTTCCTC 239
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    121 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTCTTCCTC 180

Qy    240 TCACACCT-GGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGT 298
          |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TCACACCTGGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGT 240

Qy    299 GAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCT 358
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    241 GAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCG-ATGATGCAGACCTTTCT 299

Qy    359 GTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTA 418
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    300 GTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCGGGTGGTGATGTCCTATGATCTGTA 359

Qy    419 CGT-GGCCATCTGCCACCCCCTCCGATATTT-GGCCATCATGACCTGGAGAGTCTGCATC 476
          ||| |||||||||| |||||||||||||||| |||||||||| | || ||| |   |
Db    360 CGTNGGCCATCTGNCACCCCCTCCGATATTTNGGCCATCATGAACNGGGGAGAGTCTGNC 419

Qy    477 ACC 479
          | |
Db    420 ATC 422
```

RESULT 14
BM670360
LOCUS BM670360 623 bp mRNA linear EST 27-FEB-2002
DEFINITION UI-E-DW1-ahc-b-11-0-UI.s1 UI-E-DW1 Homo sapiens cDNA clone

```
          UI-E-DW1-ahc-b-11-0-UI 3', mRNA sequence.
ACCESSION   BM670360
VERSION     BM670360.1  GI:18980257
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 623)
  AUTHORS   Bonaldo,M.F.; Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT     Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: msoares@blue.weeg.uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
             The following repetitive elements were found in this cDNA
            sequence: 95-137, >AT_rich#Low_complexity (matched compliment)
            152-189, >(CAAA)n#Simple_repeat (matched compliment)
            Seq primer: M13 Forward
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .623
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="UI-E-DW1-ahc-b-11-0-UI"
                     /clone_lib="UI-E-DW1"
                     /tissue_type="lens"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-DW1 is a normalized cDNA library containing the
                     following tissue(s): lens. The library was constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is CGATTAGCGA. This library
                     was created for the program, Gene Discovery in the Visual
                     System, supported by National Eye Institute (NEI).
                     TAG_LIB=UI-E-DW1
                     TAG_TISSUE=human lens
                     TAG_SEQ=CGATTAGCGA"
BASE COUNT      148 a    159 c    107 g    209 t
ORIGIN

  Query Match             34.1%;  Score 345.6;  DB 13;  Length 623;
  Best Local Similarity   86.7%;  Pred. No. 1.8e-83;
  Matches  392;  Conservative    0;  Mismatches   59;  Indels    1;  Gaps    1;

Qy      31 CTTGATTTCCTCTGTCATAGGGATATGGGGGACAATATAACATCCATCACAGAGTTCCTC 90
           ||| ||| | | |  ||| || |||| | | |||   |||    ||||||||||||||
Db     172 CTTTGTTTGTTTTTTAGTAGTGAAATGGTGAAAAATCAGACAATGGTCACAGAGTTCCTC 231

Qy      91 CTACTGGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 150
           |||||||||||||| | | ||||||||||||||||||||||||||||||||||||||||
Db     232 CTACTGGGATTTCTCCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 291

Qy     151 TTCTACGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCC 210
           ||||| |||||||||||||||||||| |||||||| ||||||||||||||||||||||||
Db     292 TTCTATGTCTTCACCCTGCTGGGGAATGGGACCATCCTGGGGCTCATCTCACTGGACTCC 351

Qy     211 AGACTGCACGCCCCC-TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTAC 269
           ||||| ||| ||||| |||||||||||||||||||||||| |||||| ||||||||||
Db     352 AGACTCCACACCCCCATGTACTTCTTCCTCTCACACCTGGCCGTCGTCAACATCGCCTAT 411

Qy     270 GCCTGCAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCC 329
           |||||||||||| ||||||| ||||||||||||||||||||||||||||||||||||||
Db     412 GCCTGCAACACAGTGCCCCAGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCC 471

Qy     330 TTTGCGGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTC 389
           ||||| ||| ||||||   |||||||||| |||| | ||||||    || ||||| |||
Db     472 TTTGCTGGCTGCATGACATAGACCTTTCTCTTTTTGAGTTTTGCACATACTGAATGCCTC 531

Qy     390 CTCCTGGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTG 449
           ||  ||||| |||||||||| |||| |||||||||||||||||||||| ||||||||||
Db     532 CTGTTGGTGCTGATGTCCTACGATCGGTACGTGGCCATCTGCCACCCTCTCCGATATTTC 591

Qy     450 GCCATCATGACCTGGAGAGTCTGCATCACCCT 481
           |||||||||||||| |||||||||||| ||
Db     592 ATCATCATGACCTGGAAAGTCTGCATCACTCT 623


RESULT 15
AI285267/c
LOCUS       AI285267                361 bp    mRNA    linear   EST 29-JAN-1999
DEFINITION  ql38c08.x1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1874606 3'
            similar to SW:OLF7_MOUSE P34984 OLFACTORY RECEPTOR-LIKE PROTEIN K7
            ;, mRNA sequence.
ACCESSION   AI285267
VERSION     AI285267.1  GI:3923500
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 361)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 659   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 26.
FEATURES             Location/Qualifiers
     source          1. .361
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1874606"
                     /clone_lib="NCI_CGAP_Co8"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; 1st strand cDNA was prepared from
                     colon adenocarcinoma, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector.  Library is normalized. Library was constructed by
                     Bento Soares and M. Fatima Bonaldo. "
```

```
BASE COUNT      115 a     73 c    110 g     63 t
ORIGIN

  Query Match             33.3%;  Score 337.4;  DB 9;  Length 361;
  Best Local Similarity   96.9%;  Pred. No. 2.5e-81;
  Matches  344;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy     231 TTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGG 290
           || | ||||||||||||||||| || ||||| ||||||||||||||||||||||||||||
Db     361 TTTTCCCTCTCACACCTGGCGGTTGTTGACATTGCCTACGCCTGCAACACGGTGCCCCGG 302

Qy     291 ATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAG 350
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAG 242

Qy     351 ACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTAT 410
           |||||||||||||||||||||||||||||||||||| |||||||||||||| |||| ||
Db     241 ACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTTTCCTCCTGGTGGTGAAGTCCAAT 182

Qy     411 GATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTC 470
           ||| ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db     181 GATTTGTACGTGGCCATCTGCCACCCCTCCCGATATTTGGCCATCATGACCTGGAGAGTC 122

Qy     471 TGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTT 530
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTT 62

Qy     531 GTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTT 585
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTT 7


Search completed: May  2, 2003, 06:14:26
Job time : 1553 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 04:33:51 ; Search time 2809 Seconds
(without alignments)
10505.605 Million cell updates/sec

Title: US-09-898-586-23
Perfect score: 1014
Sequence: 1 taaacacttctcctaaacca.....tatggcattgtgactgacaa 1014

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		1014	100.0	1014	6	AX197458	AX197458 Sequence
2	c	990	97.6	130030	9	AC004889	AC004889 Homo sapi
3		987.4	97.4	1012	6	AX197456	AX197456 Sequence
4		983.6	97.0	1040	6	AX197438	AX197438 Sequence
5		983.6	97.0	1040	6	AX197469	AX197469 Sequence
6	c	983.6	97.0	59130	9	AL135904	AL135904 Human DNA
7		983.6	97.0	140915	9	AC005587	AC005587 Homo sapi
8		936	92.3	974	6	AX078373	AX078373 Sequence
9		905.4	89.3	930	6	AX242191	AX242191 Sequence
10		903.6	89.1	933	6	AX451262	AX451262 Sequence
11		900.6	88.8	930	6	AX241442	AX241442 Sequence
12		900.6	88.8	930	6	AX241473	AX241473 Sequence
13		675	66.6	104955	9	AC091768	AC091768 Homo sapi
14		645.4	63.6	204316	2	AC091746	AC091746 Mus muscu
15		637.4	62.9	650	9	AF399598	AF399598 Homo sapi
16		622	61.3	933	10	AY073444	AY073444 Mus muscu
17		573.2	56.5	933	10	AY073100	AY073100 Mus muscu
18		549.8	54.2	1262	6	AX468421	AX468421 Sequence
19		546.6	53.9	66610	9	AC074386	AC074386 Homo sapi
20		545	53.7	1315	9	AB065692	AB065692 Homo sapi
21		535.4	52.8	930	6	AX242188	AX242188 Sequence
22		519.8	51.3	813	6	AX451240	AX451240 Sequence
23		518.4	51.1	1792	6	AX078369	AX078369 Sequence
24		517.8	51.1	993	6	AX242190	AX242190 Sequence
25		509.8	50.3	542	6	AX451241	AX451241 Sequence
26		508.8	50.2	542	6	AX451242	AX451242 Sequence
27		481.2	47.5	1713	6	AR092423	AR092423 Sequence
28		458	45.2	106981	2	AC120563	AC120563 Rattus no
29	c	444.4	43.8	162320	2	AC076959	AC076959 Homo sapi
30		442.2	43.6	667	10	MUSODORECA	L14566 Mouse (K7)
31		441.8	43.6	966	10	AY073724	AY073724 Mus muscu
32		440.2	43.4	966	10	AY073723	AY073723 Mus muscu
33		440.2	43.4	147341	2	AC106354	AC106354 Rattus no
34	c	437.8	43.2	10985	2	AC090838	AC090838 Homo sapi
35		430	42.4	933	10	AY073820	AY073820 Mus muscu
36		429.6	42.4	576	6	AX242035	AX242035 Sequence
37		426.8	42.1	1357	9	AB065695	AB065695 Homo sapi
38		423.8	41.8	942	10	AY073443	AY073443 Mus muscu
39		423.6	41.8	162320	2	AC076959	AC076959 Homo sapi
40		410.4	40.5	424	6	AX451249	AX451249 Sequence
41		409.6	40.4	1334	9	AB065480	AB065480 Homo sapi
42		409.6	40.4	130030	9	AC004889	AC004889 Homo sapi
43		409.2	40.4	942	6	AX241995	AX241995 Sequence
44		409.2	40.4	1207	9	AB065691	AB065691 Homo sapi
45		407.6	40.2	1333	9	AB065693	AB065693 Homo sapi

ALIGNMENTS

RESULT 1
AX197458
LOCUS AX197458 1014 bp DNA linear PAT 29-AUG-2001
DEFINITION Sequence 23 from Patent WO0151632.
ACCESSION AX197458
VERSION AX197458.1 GI:15387838
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1014)
AUTHORS Padigaru,M., Prayaga,S.K., Taupier,R.J., Mishra,V., Tchernev,V.T.,
Spytek,K.A. and Li,L.
TITLE Odorant receptor polypeptides and nucleic acids encoding same

```
  JOURNAL   Patent: WO 0151632-A 23 19-JUL-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .1014
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      218 a    281 c    216 g    299 t
ORIGIN

  Query Match           100.0%;  Score 1014;  DB 6;  Length 1014;
  Best Local Similarity  100.0%;  Pred. No. 3.2e-239;
  Matches 1014;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGG 60

Qy      61 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 120

Qy     121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 180

Qy     181 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCTGTACTTCTTCCTCT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCTGTACTTCTTCCTCT 240

Qy     241 CACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCCGGATGCTGGTGA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCCGGATGCTGGTGA 300

Qy     301 ACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTGT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTGT 360

Qy     361 TTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTACG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTACG 420

Qy     421 TGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACCC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACCC 480

Qy     481 TCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTC 540

Qy     541 TACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTTGTGAAATCTTGGCT 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 TACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTTGTGAAATCTTGGCT 600

Qy     601 GTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGCA 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGCA 660

Qy     661 ATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTGT 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 ATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTGT 720

Qy     721 GCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTCC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTCC 780

Qy     781 CACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAGA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAGA 840

Qy     841 TATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCCC 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCCC 900

Qy     901 ATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAGA 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAGA 960

Qy     961 GTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACAA 1014
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACAA 1014


RESULT 2
AC004889/c
LOCUS       AC004889              130030 bp    DNA     linear   PRI 14-JAN-1999
DEFINITION  Homo sapiens clone DJ0798C17, complete sequence.
ACCESSION   AC004889
VERSION     AC004889.1  GI:4156187
KEYWORDS    HTG.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 130030)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 130030)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   3  (bases 1 to 130030)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JAN-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Jan 14, 1999 this sequence version replaced gi:3213108.
FEATURES             Location/Qualifiers
     source          1. .130030
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DJ0798C17"
BASE COUNT    38117 a   25410 c   26567 g   39936 t
ORIGIN

  Query Match           97.6%;  Score 990;  DB 9;  Length 130030;
  Best Local Similarity   99.8%;  Pred. No. 3.4e-233;
  Matches 1012;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    2;

Qy       1 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  111297 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGG 111238

Qy      61 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  111237 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 111178

Qy     121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  111177 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 111118

Qy     181 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTCTTCCTC 239
           ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  111117 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTCTTCCTC 111058

Qy     240 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCCGGATGCTGGTG 299
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  111057 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCCGGATGCTGGTG 110998

Qy     300 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 359
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  110997 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 110938
```

```
Qy    360 TTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 110937 TTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 110878

Qy    420 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 110877 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 110818

Qy    480 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 110817 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 110758

Qy    540 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTTGTGAAATCTTGGC 599
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db 110757 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAATCTTGGC 110699

Qy    600 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 110698 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 110639

Qy    660 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 110638 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 110579

Qy    720 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTC 779
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 110578 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTC 110519

Qy    780 CCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 839
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 110518 CCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 110459

Qy    840 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 110458 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 110399

Qy    900 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 110398 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 110339

Qy    960 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1013
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 110338 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 110285
```

RESULT 3
AX197456
LOCUS AX197456 1012 bp DNA linear PAT 29-AUG-2001
DEFINITION Sequence 21 from Patent WO0151632.
ACCESSION AX197456
VERSION AX197456.1 GI:15387837
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1012)
AUTHORS Padigaru,M., Prayaga,S.K., Taupier,R.J., Mishra,V., Tchernev,V.T., Spytek,K.A. and Li,L.
TITLE Odorant receptor polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 0151632-A 21 19-JUL-2001;
Curagen Corporation (US)
FEATURES Location/Qualifiers
source 1. .1012
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 219 a 281 c 215 g 297 t
ORIGIN

Query Match 97.4%; Score 987.4; DB 6; Length 1012;
Best Local Similarity 99.7%; Pred. No. 1.2e-232;
Matches 1010; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

```
Qy      2 AAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGGG 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGAG 60

Qy     62 ACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATTC 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATTC 120

Qy    122 AGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGGA 181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGGA 180

Qy    182 CCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTCTTCCTCT 240
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    181 CCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTCTTCCTCT 240

Qy    241 CACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGA 300

Qy    301 ACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTGT 360

Qy    361 TTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTACG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTACG 420

Qy    421 TGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACCC 480

Qy    481 TCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTC 540

Qy    541 TACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTTGTGAAATCTTGGCT 600
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    541 TACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAATCTTGGCT 599

Qy    601 GTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGCA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 GTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGCA 659

Qy    661 ATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTGT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    660 ATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTGT 719

Qy    721 GCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTCC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    720 GCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTCC 779

Qy    781 CACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    780 CACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAGA 839

Qy    841 TATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCCC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    840 TATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCCC 899

Qy    901 ATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAGA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    900 ATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAGA 959

Qy    961 GTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1013
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    960 GTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1012
```

RESULT 4
AX197438
LOCUS AX197438 1040 bp DNA linear PAT 29-AUG-2001
DEFINITION Sequence 3 from Patent WO0151632.

```
ACCESSION   AX197438
VERSION     AX197438.1  GI:15387828
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1040)
  AUTHORS   Padigaru,M., Prayaga,S.K., Taupier,R.J., Mishra,V., Tchernev,V.T.,
            Spytek,K.A. and Li,L.
  TITLE     Odorant receptor polypeptides and nucleic acids encoding same
  JOURNAL   Patent: WO 0151632-A 3 19-JUL-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .1040
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      229 a    285 c    221 g    305 t
ORIGIN

  Query Match             97.0%;  Score 983.6;  DB 6;  Length 1040;
  Best Local Similarity   99.4%;  Pred. No. 1e-231;
  Matches 1008;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy       1 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      28 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGA 87

Qy      61 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 120
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      88 GACAATATAACATCCATCAGAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 147

Qy     121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     148 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 207

Qy     181 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTCTTCCTC 239
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     208 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTCTTCCTC 267

Qy     240 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTG 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     268 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTG 327

Qy     300 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     328 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 387

Qy     360 TTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     388 TTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 447

Qy     420 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     448 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 507

Qy     480 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     508 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 567

Qy     540 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTTGTGAAATCTTGGC 599
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     568 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAATCTTGGC 626

Qy     600 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     627 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 686

Qy     660 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     687 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 746

Qy     720 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTC 779
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     747 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCCGCACCTGCTTCTC 806

Qy     780 CCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 839
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     807 CCACCTCTGTGTGATTGGACTCGTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 866

Qy     840 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     867 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 926

Qy     900 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     927 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 986

Qy     960 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1013
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     987 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1040


RESULT 5
AX197469
LOCUS       AX197469                1040 bp    DNA     linear   PAT 29-AUG-2001
DEFINITION  Sequence 34 from Patent WO0151632.
ACCESSION   AX197469
VERSION     AX197469.1  GI:15387841
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1040)
  AUTHORS   Padigaru,M., Prayaga,S.K., Taupier,R.J., Mishra,V., Tchernev,V.T.,
            Spytek,K.A. and Li,L.
  TITLE     Odorant receptor polypeptides and nucleic acids encoding same
  JOURNAL   Patent: WO 0151632-A 34 19-JUL-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .1040
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      229 a    285 c    221 g    305 t
ORIGIN

  Query Match             97.0%;  Score 983.6;  DB 6;  Length 1040;
  Best Local Similarity   99.4%;  Pred. No. 1e-231;
  Matches 1008;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy       1 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      28 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGA 87

Qy      61 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 120
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      88 GACAATATAACATCCATCAGAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 147

Qy     121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     148 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 207

Qy     181 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTCTTCCTC 239
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     208 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTCTTCCTC 267

Qy     240 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTG 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     268 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTG 327

Qy     300 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     328 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 387

Qy     360 TTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 419
```

```
Db    388 TTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 447

Qy    420 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 479

Db    448 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 507

Qy    480 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 539

Db    508 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 567

Qy    540 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAAATCTTGGC 599

Db    568 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAATCTTGGC 626

Qy    600 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 659

Db    627 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 686

Qy    660 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 719

Db    687 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 746

Qy    720 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTC 779

Db    747 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCCGCACCTGCTTCTC 806

Qy    780 CCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 839

Db    807 CCACCTCTGTGTGATTGGACTCGTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 866

Qy    840 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 899

Db    867 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 926

Qy    900 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 959

Db    927 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 986

Qy    960 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1013

Db    987 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1040


RESULT 6
AL135904/c
LOCUS       AL135904               59130 bp    DNA     linear   PRI 20-JUL-2000
DEFINITION  Human DNA sequence from clone RP5-1005H11 on chromosome 6 Contains
            part of the gene for a 7 transmembrane receptor (rhodopsin family)
            (olfactory receptor like) protein, the gene for WUGSC:H_DJ0988G15.3
            protein, part of the PDNP3 (phosphodiesterase I/nucleotide
            pyrophosphatase 3) gene, ESTs, STSs and GSSs, complete sequence.
ACCESSION   AL135904
VERSION     AL135904.11  GI:7159399
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 59130)
  AUTHORS   Lovell,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUL-2000) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Mar 6, 2000 this sequence version replaced gi:7105768.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence has been finished according to sequence map criteria
            as follows. An attempt is made to resolve all sequencing problems,
            such as compressions and repeats, but not necessarily within known
            annotated human repeat sequence elements (e.g. Alu). Where the
            sequence is ambiguous, there is an annotation using the 'unsure'
            feature key.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP5-1005H11 is from the library RPCI-5 constructed at the Roswell
            Park Cancer Institute by the group of Pieter de Jong. For further
            details see http://bacpac.med.buffalo.edu/
            VECTOR: pCYPAC2
            IMPORTANT: This sequence is not the entire insert of clone
            RP5-1005H11 It may be shorter because we sequence overlapping
            sections only once, except for a 100 base overlap.
            The true right end of clone RP5-1005H11 is at 59130 in this
            sequence. The true right end of clone RP5-914N13 is at 105 in this
            sequence.
FEATURES             Location/Qualifiers
     source          1. .59130
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP5-1005H11"
                     /clone_lib="RPCI-5"
     repeat_region   1. .68
                     /note="AluS repeat: matches 1. .68 of consensus"
     repeat_region   546. .844
                     /note="AluJo repeat: matches 2. .301 of consensus"
     repeat_region   870. .940
                     /note="U4 repeat: matches 1. .72 of consensus"
     repeat_region   1691. .2906
                     /note="L2 repeat: matches 1425. .2737 of consensus"
     repeat_region   3457. .3960
                     /note="HAL1 repeat: matches 420. .971 of consensus"
     repeat_region   3971. .4106
                     /note="MLT1D repeat: matches 1. .129 of consensus"
     repeat_region   4249. .4282
                     /note="17 copies 2 mer tg 100% conserved"
     repeat_region   4966. .5147
                     /note="AluJb repeat: matches 1. .182 of consensus"
     repeat_region   5216. .5307
                     /note="L2 repeat: matches 2405. .2503 of consensus"
     repeat_region   5495. .5794
                     /note="AluSx repeat: matches 1. .296 of consensus"
     repeat_region   5891. .6180
                     /note="AluSx repeat: matches 1. .291 of consensus"
     repeat_region   7607. .7618
                     /note="L1M1 repeat: matches 4856. .4867 of consensus"
     repeat_region   7619. .8347
                     /note="L1PA16 repeat: matches 5421. .6156 of consensus"
     repeat_region   8348. .8898
                     /note="L1M1 repeat: matches 4866. .5403 of consensus"
     repeat_region   9018. .9866
                     /note="TIGGER1 repeat: matches 1. .836 of consensus"
     repeat_region   9867. .10189
                     /note="AluSx repeat: matches 1. .309 of consensus"
     repeat_region   10190. .11723
                     /note="TIGGER1 repeat: matches 836. .2418 of consensus"
     repeat_region   12404. .12715
                     /note="AluSx repeat: matches 1. .312 of consensus"
     misc_feature    complement(13947. .14176)
                     /note="match: GSS: Em:B74942"
     repeat_region   14280. .14383
                     /note="AluSg/x repeat: matches 190. .294 of consensus"
     repeat_region   15340. .15508
                     /note="L2 repeat: matches 2563. .2750 of consensus"
```

```
repeat_region   15684. .15968
                /note="AluSx repeat: matches 35. .312 of consensus"
repeat_region   16527. .16612
                /note="MIR repeat: matches 50. .138 of consensus"
repeat_region   16731. .17027
                /note="AluSg repeat: matches 1. .295 of consensus"
repeat_region   17609. .17747
                /note="MER91A repeat: matches 1. .153 of consensus"
repeat_region   17867. .18162
                /note="AluY repeat: matches 1. .299 of consensus"
repeat_region   18461. .18950
                /note="MER1A repeat: matches 2. .527 of consensus"
repeat_region   19185. .19224
                /note="20 copies 2 mer tg 85% conserved"
misc_feature    19474. .19712
                /note="match: STS: Em:G16149"
repeat_region   20037. .20361
                /note="AluSx repeat: matches 1. .311 of consensus"
gene            complement(21566. .22384)
                /gene="dJ1005H11.1"
CDS             complement(21566. .>22384)
                /gene="dJ1005H11.1"
                /note="match: proteins: Sw:P47881 Tr:O70265 Tr:O70266
                Tr:O70267 Tr:O70268 Tr:Q62942 Tr:O95918 Tr:O95047"
                /codon_start=1
                /evidence=not_experimental
                /product="dJ1005H11.1 (7 TRANSMEMBRANE RECEPTOR (RHODOPSIN
                FAMILY) (OLFACTORY RECEPTOR LIKE) PROTEIN))"
                /protein_id="CAB99212.1"
                /db_xref="GI:9368991"
                /translation="LGNGTILGLISLDSRLHAPMYFFLSHLAVVDIAYACNTVPRMLV
                NLLHPAKPISFAGRMMQTFLFSTFAVTECLLLVVMSYDLYVAICHPLRYLAIMTWRVC
                ITLAVTSWTTGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILAVLKLACADTHINENMV
                LAGAISGLVGPLSTIVVSYMCILCAILQIQSREVQRKAFRTCFSHLCVIGLVYGTAII
                MYVGPRYGNPKEQKKYLLLFHSLFNPMLNPLICSLRNSEVKNTLKRVLGVERAL"
misc_feature    complement(22229. .22521)
                /note="match: STS: Em:H39853"
repeat_region   22385. .22489
                /note="L2 repeat: matches 2461. .2551 of consensus"
repeat_region   22696. .23011
                /note="AluSx repeat: matches 1. .312 of consensus"
misc_feature    23044. .23513
                /note="match: STS: Em:AI066560"
misc_feature    complement(23123. .23716)
                /note="match: GSS: Em:AQ276140"
misc_feature    complement(23300. .23543)
                /note="match: GSS: Em:AQ058777"
misc_feature    23751. .24305
                /note="match: GSS: Em:AQ798220"
misc_feature    23753. .24176
                /note="match: GSS: Em:AQ550610"
misc_feature    23762. .24465
                /note="match: GSS: Em:B86674"
repeat_region   24220. .24343
                /note="FLAM_C repeat: matches 1. .124 of consensus"
misc_feature    complement(24681. .25395)
                /note="match: GSS: Em:AQ627093"
misc_feature    complement(24780. .25383)
                /note="match: GSS: Em:AQ627184"
repeat_region   26053. .26548
                /note="TIGGER1 repeat: matches 1. .521 of consensus"
repeat_region   26549. .26861
                /note="L1M4 repeat: matches 4025. .4339 of consensus"
repeat_region   26865. .27121
                /note="Tigger3b repeat: matches 2. .272 of consensus"
repeat_region   27122. .27424
                /note="AluSx repeat: matches 1. .302 of consensus"
repeat_region   27425. .27579
                /note="Tigger3b repeat: matches 272. .451 of consensus"
repeat_region   27582. .27658
                /note="Tigger3b repeat: matches 1155. .1231 of consensus"
repeat_region   27659. .29067
                /note="L1M4 repeat: matches 2585. .4035 of consensus"
repeat_region   29189. .29287
                /note="L1M4 repeat: matches 2499. .2600 of consensus"
misc_feature    29316. .29499
                /note="match: STS: Em:AA931824"
misc_feature    29507. .29990
                /note="match: STS: Em:AA931824"
misc_feature    29530. .29981
                /note="match: GSS: Em:AQ356766"
misc_feature    29530. .29737
                /note="match: GSS: Em:AQ253729"
misc_feature    29533. .29855
                /note="match: GSS: Em:AQ319835"
misc_feature    29538. .29943
                /note="match: STS: Em:AA235111"
misc_feature    29560. .29787
                /note="match: STS: Em:G27651 Em:T99793"
gene            complement(29779. .32112)
                /gene="dJ1005H11.2"
CDS             complement(29779. .32112)
                /gene="dJ1005H11.2"
                /note="match: proteins: Tr:O95046"
                /codon_start=1
                /evidence=not_experimental
                /product="dJ1005H11.2 (WUGSC:H_DJ0988G15.3 PROTEIN)"
                /protein_id="CAB99213.1"
                /db_xref="GI:9368992"
                /translation="MEEPGATPQPYLGLVLEELGRVVAALPESMRPDENPYGFPSELV
```

```
Query Match             97.0%;  Score 983.6;  DB 9;  Length 59130;
Best Local Similarity   99.4%;  Pred. No. 1.3e-231;
Matches 1008;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy       1 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22552 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGA 22493

Qy      61 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 120
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   22492 GACAATATAACATCCATCAGAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 22433

Qy     121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22432 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 22373

Qy     181 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTCTTCCTC 239
           |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   22372 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTCTTCCTC 22313

Qy     240 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTG 299
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22312 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTG 22253

Qy     300 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 359
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22252 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 22193

Qy     360 TTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 419
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22192 TTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 22133

Qy     420 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 479
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22132 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 22073

Qy     480 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 539
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22072 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 22013

Qy     540 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAAATCTTGGC 599
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   22012 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAATCTTGGC 21954

Qy     600 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 659
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21953 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 21894

Qy    660 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 719
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21893 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 21834

Qy    720 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTC 779
                |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  21833 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCCGCACCTGCTTCTC 21774

Qy    780 CCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 839
                |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  21773 CCACCTCTGTGTGATTGGACTCGTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 21714

Qy    840 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 899
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21713 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 21654

Qy    900 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 959
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21653 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 21594

Qy    960 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1013
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21593 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 21540


RESULT 7
AC005587
LOCUS       AC005587               140915 bp    DNA     linear   PRI 21-DEC-1999
DEFINITION  Homo sapiens PAC clone RP5-988G15 from 7q33-q35, complete sequence.
ACCESSION   AC005587
VERSION     AC005587.1  GI:4156166
KEYWORDS    HTG.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 140915)
  AUTHORS   Hou,S., Wohldmann,P. and Le,T.
  TITLE     The sequence of Homo sapiens PAC clone RP5-988G15
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 140915)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-SEP-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   3  (bases 1 to 140915)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JAN-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 140915)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-1999) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   5  (bases 1 to 140915)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-DEC-1999) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Jan 14, 1999 this sequence version replaced gi:3907499.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_DJ0988G15
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded or sequenced with an alternate
            chemistry; an attempt was made to resolve all sequencing problems,
            such as compressions and repeats; all regions were covered by
            sequence from more than one subclone; and the assembly was
            confirmed by restriction digest.

            MAPPING INFORMATION:
            The sequence of this clone was established as part of a mapping and
            sequencing collaboration between the NHGRI Chromosome 7 Mapping
            Project (Eric D. Green, Director), John D. McPherson in the
            Department of Genetics (Washington University), and the Washington
            University Genome Sequencing Center.  For additional information
            about the map position of this sequence, see
            http://www.nhgri.nih.gov/DIR/GTB/CHR7 , send
            mailto:egreen@nhgri.nih.gov , or see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            This clone was derived from human PAC library RPCI-5, prepared by
            Pieter de Jong and coworkers at the Roswell Park Cancer Institute
            (http://bacpac.med.buffalo.edu) using the method described by
            Ioannou et al., Nature Genetics 6:84-9 (1994).  The library is from
            one male donor.
            The clone may be obtained either from Genome Systems, Inc.
            (http://www.genomesystems.com) or Research Genetics, Inc.
            (http://www.resgen.com); or from Pieter de Jong.
            VECTOR:  pCYPAC2
            NEIGHBORING SEQUENCE INFORMATION:
            Actual start of this clone is at base position 1 of RP5-988G15;
            actual end is at 140915 of RP5-988G15.
FEATURES             Location/Qualifiers
     source          1. .140915
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q33-q35"
                     /clone="RP5-988G15"
                     /clone_lib="RPCI-5"
     repeat_region   13. .2090
                     /rpt_family="L1"
     repeat_region   2097. .2299
                     /rpt_family="L1"
     repeat_region   2300. .2603
                     /rpt_family="Alu"
     repeat_region   2604. .4120
                     /rpt_family="L1"
     repeat_region   4123. .4166
                     /rpt_family="(TA)n"
     repeat_region   4168. .4307
                     /rpt_family="L1"
     repeat_region   6635. .6668
                     /rpt_family="AT_rich"
     repeat_region   6829. .7260
                     /rpt_family="L1"
     repeat_region   7261. .7642
                     /rpt_family="MaLR"
     repeat_region   7643. .7923
                     /rpt_family="L1"
     repeat_region   7982. .8283
                     /rpt_family="Alu"
     repeat_region   8286. .8412
                     /rpt_family="L1"
     repeat_region   8413. .8708
                     /rpt_family="Alu"
     repeat_region   8709. .8838
                     /rpt_family="L1"
```

```
repeat_region   8840. .8886
                /rpt_family="MaLR"
repeat_region   9111. .9139
                /rpt_family="AT_rich"
repeat_region   9204. .9490
                /rpt_family="Alu"
repeat_region   9646. .9924
                /rpt_family="Alu"
repeat_region   10286. .10455
                /rpt_family="L2"
repeat_region   10759. .10838
                /rpt_family="Mariner"
repeat_region   11127. .11266
                /rpt_family="L2"
repeat_region   11850. .12154
                /rpt_family="Alu"
repeat_region   12172. .12211
                /rpt_family="AT_rich"
repeat_region   12352. .12469
                /rpt_family="(TACG)n"
repeat_region   12458. .12607
                /rpt_family="(CATA)n"
repeat_region   12580. .12733
                /rpt_family="(TA)n"
repeat_region   12660. .12777
                /rpt_family="(CATA)n"
repeat_region   12778. .13031
                /rpt_family="Alu"
repeat_region   13226. .13535
                /rpt_family="L2"
repeat_region   15279. .15520
                /rpt_family="MIR"
repeat_region   15954. .16022
                /rpt_family="MaLR"
repeat_region   16023. .16332
                /rpt_family="MER2_type"
repeat_region   16333. .16371
                /rpt_family="MaLR"
repeat_region   16630. .16742
                /rpt_family="MaLR"
repeat_region   16816. .17068
                /rpt_family="L1"
repeat_region   17088. .17348
                /rpt_family="L1"
repeat_region   17857. .17979
                /rpt_family="Alu"
repeat_region   18014. .18506
                /rpt_family="MaLR"
repeat_region   18707. .19081
                /rpt_family="MaLR"
repeat_region   19135. .20717
                /rpt_family="Retroviral"
repeat_region   20728. .20748
                /rpt_family="AT_rich"
repeat_region   21444. .21599
                /rpt_family="MIR"
misc_feature    21688. .22279
                /note="60S ribosomal protein L15 pseudogene"
repeat_region   22506. .22540
                /rpt_family="AT_rich"
repeat_region   22557. .22583
                /rpt_family="AT_rich"
repeat_region   23139. .23162
                /rpt_family="(TAA)n"
repeat_region   23696. .24012
                /rpt_family="Alu"
repeat_region   24088. .24128
                /rpt_family="AT_rich"
repeat_region   24827. .24849
                /rpt_family="AT_rich"
repeat_region   24850. .25147
                /rpt_family="Alu"
repeat_region   25148. .25173
                /rpt_family="AT_rich"
repeat_region   25951. .25974
                /rpt_family="(CAAAA)n"
repeat_region   26066. .26144
                /rpt_family="Alu"
repeat_region   26388. .26653
                /rpt_family="L1"
repeat_region   27148. .27600
                /rpt_family="L2"
repeat_region   28002. .28294
                /rpt_family="Alu"
```

Query Match 97.0%; Score 983.6; DB 9; Length 140915;
Best Local Similarity 99.4%; Pred. No. 1.3e-231;
Matches 1008; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

```
Qy       1 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  101545 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGA 101604

Qy      61 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 120
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  101605 GACAATATAACATCCATCAGAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 101664

Qy     121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  101665 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 101724

Qy     181 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTCTTCCTC 239
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  101725 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTCTTCCTC 101784

Qy     240 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTG 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  101785 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTG 101844

Qy     300 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  101845 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 101904

Qy     360 TTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  101905 TTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 101964

Qy     420 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  101965 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 102024

Qy     480 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  102025 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 102084

Qy     540 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAAATCTTGGC 599
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db  102085 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAATCTTGGC 102143

Qy     600 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  102144 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 102203

Qy     660 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  102204 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 102263

Qy     720 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTC 779
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  102264 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCCGCACCTGCTTCTC 102323

Qy     780 CCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 839
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db  102324 CCACCTCTGTGTGATTGGACTCGTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 102383

Qy     840 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 899
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 102384 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 102443

Qy    900 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 102444 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 102503

Qy    960 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1013
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 102504 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 102557


RESULT 8
AX078373
LOCUS       AX078373                974 bp    DNA     linear   PAT 22-FEB-2001
DEFINITION  Sequence 41 from Patent WO0107612.
ACCESSION   AX078373
VERSION     AX078373.1  GI:13158042
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 974)
  AUTHORS   Au-Young,J., Bandman,O., Tang,Y.T., Yue,H., Azimzai,Y., Burford,N.,
            Baughn,M.R., Lu,D.A., Hillman,J.L., Patterson,C. and Lal,P.
  TITLE     Receptors and associated proteins
  JOURNAL   Patent: WO 0107612-A 41 01-FEB-2001;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .974
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 2279267CB1"
BASE COUNT      206 a    270 c    214 g    284 t
ORIGIN

  Query Match           92.3%;  Score 936;  DB 6;  Length 974;
  Best Local Similarity  99.8%;  Pred. No. 5.5e-220;
  Matches  958;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    2;

Qy      55 ATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCA 114
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCA 60

Qy     115 AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGG 174
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGG 120

Qy     175 AACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTC 233
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db     121 AACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTC 180

Qy     234 TTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATG 293
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATG 240

Qy     294 CTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 353
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 300

Qy     354 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGAT 413
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGAT 360

Qy     414 CTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 473
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 420

Qy     474 ATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTG 533
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTG 480

Qy     534 TTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAAAT 593
           |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     481 TTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAAT 539

Qy     594 CTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGC 653
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     540 CTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGC 599

Qy     654 CGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCAT 713
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 CGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCAT 659

Qy     714 CCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTG 773
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     660 CCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTG 719

Qy     774 CTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGG 833
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     720 CTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGG 779

Qy     834 ACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTT 893
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     780 ACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTT 839

Qy     894 TAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTT 953
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     840 TAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTT 899

Qy     954 GAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1013
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     900 GAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 959


RESULT 9
AX242191
LOCUS       AX242191                930 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 939 from Patent WO0127158.
ACCESSION   AX242191
VERSION     AX242191.1  GI:15799066
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 930)
  AUTHORS   Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and
            Yanai,I.
  TITLE     Olfactory receptor sequences
  JOURNAL   Patent: WO 0127158-A 939 19-APR-2001;
            Digiscents (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES             Location/Qualifiers
     source          1. .930
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="(H38g789 nucleotide)"
BASE COUNT      192 a    266 c    201 g    271 t
ORIGIN

  Query Match           89.3%;  Score 905.4;  DB 6;  Length 930;
  Best Local Similarity  99.7%;  Pred. No. 1.9e-212;
  Matches  928;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    2;

Qy      55 ATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCA 114
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCA 60

Qy     115 AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGG 174
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGG 120

Qy     175 AACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTC 233
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db     121 AACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTC 180
```

```
Qy    234 TTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATG 293
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATG 240

Qy    294 CTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 300

Qy    354 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGAT 413
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGAT 360

Qy    414 CTGTACGTGGCCATCTGCCACCCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 473
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTGTACGTGGCCATCTGCCACCCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 420

Qy    474 ATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTG 533
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTG 480

Qy    534 TTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAAAT 593
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    481 TTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAAT 539

Qy    594 CTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGC 653
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 CTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGC 599

Qy    654 CGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCAT 713
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 CGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCAT 659

Qy    714 CCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTG 773
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    660 CCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTG 719

Qy    774 CTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGG 833
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    720 CTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGG 779

Qy    834 ACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTT 893
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    780 ACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCCGCTGTTTCACAGCCTCTT 839

Qy    894 TAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTT 953
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    840 TAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTT 899

Qy    954 GAAGAGAGTGCTGGGAGTAGAAAGGGCTTTA 984
          |||||||||||||||||||||||||||||||
Db    900 GAAGAGAGTGCTGGGAGTAGAAAGGGCTTTA 930
```

```
RESULT 10
AX451262
LOCUS       AX451262                 933 bp    DNA     linear   PAT 03-JUL-2002
DEFINITION  Sequence 1455 from Patent WO0214501.
ACCESSION   AX451262
VERSION     AX451262.1  GI:21698331
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1
  AUTHORS   Hubert,R.S., Raitano,A.B., Faris,M., Challita-Eid,P.M., Ge,W. and
            Jakobovits,A.
  TITLE     Nucleic acids and corresponding proteins entitled phor1-a11 and
            phor1-f5d6 useful in treatment and detection of cancer
  JOURNAL   Patent: WO 0214501-A 1455 21-FEB-2002;
            Agensys, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .933
                     /organism="unidentified"
                     /db_xref="taxon:32644"
                     /note="PHOR1-F5D6"
     CDS             1. .933
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD37777.1"
                     /db_xref="GI:21698332"
                     /translation="MGDNITSIREFLLLGFPVGPRIQMLLFGLFSLFYVFTLLGNGTI
                     LGLISLDSRLHAPMYFFLSHLAVVDIAYACNTVPRMLVNLLHPAKPISFAGRMMQTFL
                     FSTFAVTECLLLVVMSYDLYVAICHPLRYLAIMTWRVCITLAVTSWTTGVLLSLIHLV
                     LLLPLPFCRPQKIYHFFCEILAVLKLACADTHINENMVLAGAISGLVGPLSTIVVSYM
                     CILCAILQIQSREVQRKAFRTCFSHLCVIGLVYGTAIIMYVGPRYGNPKEQKKYLLLF
                     HSLFNPMLNPLICSLRNSEVKNTLKRVLGVERAL"
BASE COUNT      194 a    265 c    203 g    271 t
ORIGIN
```

```
  Query Match             89.1%;  Score 903.6;  DB 6;  Length 933;
  Best Local Similarity   99.4%;  Pred. No. 5.4e-212;
  Matches  928;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy     55 ATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCA 114
          ||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db      1 ATGGGAGACAATATAACATCCATCAGAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCA 60

Qy    115 AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGG 174
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGG 120

Qy    175 AACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTC 233
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db    121 AACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTC 180

Qy    234 TTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATG 293
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATG 240

Qy    294 CTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 300

Qy    354 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGAT 413
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGAT 360

Qy    414 CTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 473
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 420

Qy    474 ATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTG 533
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTG 480

Qy    534 TTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAAAT 593
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    481 TTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAAT 539

Qy    594 CTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGC 653
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 CTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGC 599

Qy    654 CGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCAT 713
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 CGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCAT 659

Qy    714 CCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTG 773
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    660 CCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCCGCACCTG 719

Qy    774 CTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGG 833
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    720 CTTCTCCCACCTCTGTGTGATTGGACTCGTTTATGGCACAGCCATTATCATGTATGTTGG 779

Qy    834 ACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTT 893
```

```
Db    780 ACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTT 839

Qy    894 TAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTT 953
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    840 TAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTT 899

Qy    954 GAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGA 987
          |||||||||||||||||||||||||||||||||
Db    900 GAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGA 933


RESULT 11
AX241442
LOCUS       AX241442                930 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 190 from Patent WO0127158.
ACCESSION   AX241442
VERSION     AX241442.1  GI:15798317
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 930)
  AUTHORS   Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and
            Yanai,I.
  TITLE     Olfactory receptor sequences
  JOURNAL   Patent: WO 0127158-A 190 19-APR-2001;
            Digiscents (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES             Location/Qualifiers
     source          1. .930
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="(H38g38 nucleotide)"
BASE COUNT      193 a    265 c    202 g    270 t
ORIGIN

  Query Match           88.8%;  Score 900.6;  DB 6;  Length 930;
  Best Local Similarity   99.4%;  Pred. No. 2.9e-211;
  Matches  925;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy     55 ATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCA 114
          ||||| |||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db      1 ATGGGAGACAATATAACATCCATCAGAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCA 60

Qy    115 AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGG 174
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGG 120

Qy    175 AACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTC 233
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db    121 AACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTC 180

Qy    234 TTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATG 293
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATG 240

Qy    294 CTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 300

Qy    354 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGAT 413
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGAT 360

Qy    414 CTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 473
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 420

Qy    474 ATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTG 533
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTG 480

Qy    534 TTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAAAT 593
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    481 TTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAAT 539

Qy    594 CTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGC 653
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 CTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGC 599

Qy    654 CGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCAT 713
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 CGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCAT 659

Qy    714 CCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTG 773
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    660 CCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCCGCACCTG 719

Qy    774 CTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGG 833
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    720 CTTCTCCCACCTCTGTGTGATTGGACTCGTTTATGGCACAGCCATTATCATGTATGTTGG 779

Qy    834 ACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTT 893
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    780 ACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTT 839

Qy    894 TAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTT 953
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    840 TAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTT 899

Qy    954 GAAGAGAGTGCTGGGAGTAGAAAGGGCTTTA 984
          |||||||||||||||||||||||||||||
Db    900 GAAGAGAGTGCTGGGAGTAGAAAGGGCTTTA 930


RESULT 12
AX241473
LOCUS       AX241473                930 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 221 from Patent WO0127158.
ACCESSION   AX241473
VERSION     AX241473.1  GI:15798348
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 930)
  AUTHORS   Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and
            Yanai,I.
  TITLE     Olfactory receptor sequences
  JOURNAL   Patent: WO 0127158-A 221 19-APR-2001;
            Digiscents (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES             Location/Qualifiers
     source          1. .930
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="(H38g70 nucleotide)"
BASE COUNT      193 a    265 c    202 g    270 t
ORIGIN

  Query Match           88.8%;  Score 900.6;  DB 6;  Length 930;
  Best Local Similarity   99.4%;  Pred. No. 2.9e-211;
  Matches  925;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy     55 ATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCA 114
          ||||| |||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db      1 ATGGGAGACAATATAACATCCATCAGAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCA 60

Qy    115 AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGG 174
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGG 120

Qy    175 AACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTC 233
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db    121 AACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTC 180

Qy    234 TTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATG 293
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATG 240

Qy    294 CTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 300

Qy    354 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGAT 413
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGAT 360

Qy    414 CTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 473
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 420

Qy    474 ATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTG 533
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTG 480

Qy    534 TTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTTGTGAAAT 593
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    481 TTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAAT 539

Qy    594 CTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGC 653
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 CTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGC 599

Qy    654 CGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCAT 713
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 CGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCAT 659

Qy    714 CCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTG 773
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    660 CCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCCGCACCTG 719

Qy    774 CTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGG 833
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    720 CTTCTCCCACCTCTGTGTGATTGGACTCGTTTATGGCACAGCCATTATCATGTATGTTGG 779

Qy    834 ACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTT 893
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    780 ACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTT 839

Qy    894 TAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTT 953
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    840 TAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTT 899

Qy    954 GAAGAGAGTGCTGGGAGTAGAAAGGGCTTTA 984
          |||||||||||||||||||||||||||||||
Db    900 GAAGAGAGTGCTGGGAGTAGAAAGGGCTTTA 930
```

```
RESULT 13
AC091768
LOCUS       AC091768               104955 bp    DNA     linear   PRI 09-JAN-2002
DEFINITION  Homo sapiens BAC clone RP11-703N5 from 7, complete sequence.
ACCESSION   AC091768
VERSION     AC091768.4  GI:15145624
KEYWORDS    HTG.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 104955)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
   PUBMED   9847074
REFERENCE   2  (bases 1 to 104955)
  AUTHORS   Harris,A., Haakenson,W. and Spalding,L.
  TITLE     The sequence of Homo sapiens BAC clone RP11-703N5
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 104955)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-MAY-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 104955)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-AUG-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 104955)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JAN-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Aug 9, 2001 this sequence version replaced gi:14550325.
            ---------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            ---------- Summary Statistics
                Center project name: H_NH0703N05
            ----------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            The sequence of this clone was established as part of a mapping and
            sequencing collaboration between the NHGRI Chromosome 7 Mapping
            Project (Eric D. Green, Director), John D. McPherson in the
            Department of Genetics (Washington University), and the Washington
            University Genome Sequencing Center.  For additional information
            about the map position of this sequence, see
            http://www.nhgri.nih.gov/DIR/GTB/CHR7 , send
            mailto:egreen@nhgri.nih.gov , or see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries.  Genomics 51:1-8.  The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at the Roswell Park Cancer Institute
            (http://bacpac.med.buffalo.edu)
            VECTOR:  pBACe3.6
            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is RP4-669B10, 2000 bp overlap; the
            clone sequenced to the right is RP4-798C17, 2000 bp overlap.
            Actual start of this clone is at base position 123947 of
            RP4-669B10; actual end is at base position 64342 of RP11-798C17.

            RP11-703N5 contains a single plasmid region from 4530 to 4616 and
            4668 to 4721.
FEATURES              Location/Qualifiers
     source           1. .104955
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /chromosome="7"
```

```
                     /map="7"
                     /clone="RP11-703N5"
                     /clone_lib="RPCI-11"
     repeat_region   1. .162
                     /rpt_family="L1"
     repeat_region   147. .169
                     /rpt_family="AT_rich"
     repeat_region   928. .1614
                     /rpt_family="L2"
     repeat_region   2091. .2204
                     /rpt_family="L2"
     repeat_region   2649. .2713
                     /rpt_family="ERV1"
     repeat_region   2651. .2710
                     /rpt_family="(CA)n"
     repeat_region   4071. .4108
                     /rpt_family="(TCCA)n"
     repeat_region   4472. .4511
                     /rpt_family="L2"
     misc_feature    5679. .5892
                     /note="similar to Homo sapiens EST AA298029
                     (NID:g1950363)"
     repeat_region   7802. .7825
                     /rpt_family="AT_rich"
     repeat_region   8020. .8173
                     /rpt_family="MIR"
     repeat_region   9639. .9865
                     /rpt_family="CR1"
     repeat_region   10327. .10490
                     /rpt_family="Alu"
     repeat_region   10476. .10498
                     /rpt_family="(CAAA)n"
     repeat_region   11042. .11088
                     /rpt_family="GA-rich"
     repeat_region   11387. .11406
                     /rpt_family="(T)n"
     repeat_region   11947. .12268
                     /rpt_family="L1"
     misc_feature    12489. .13017
                     /note="match to EST AI478847 (NID:g4372015) tm24e12.x1"
     repeat_region   12793. .12973
                     /rpt_family="L1"
     repeat_region   12974. .13192
                     /rpt_family="MaLR"
     misc_feature    13147. .13148
                     /note="match to EST AI478847 (NID:g4372015) tm24e12.x1"
     repeat_region   13193. .13485
                     /rpt_family="Alu"
     repeat_region   13486. .13668
                     /rpt_family="MaLR"
     repeat_region   13669. .14177
                     /rpt_family="MaLR"
     repeat_region   14177. .15248
                     /rpt_family="MaLR"
     repeat_region   15249. .16333
                     /rpt_family="ERVK"
     repeat_region   16334. .16622
                     /rpt_family="MaLR"
     repeat_region   16628. .16988
                     /rpt_family="MaLR"
     repeat_region   16989. .17116
                     /rpt_family="MaLR"
     repeat_region   17119. .17273
                     /rpt_family="L1"
     repeat_region   18350. .18588
                     /rpt_family="ERVL"
     repeat_region   18658. .18733
                     /rpt_family="ERVL"
     repeat_region   19319. .19982
                     /rpt_family="L1"
     repeat_region   20126. .20161
                     /rpt_family="AT_rich"
     repeat_region   20389. .21008
                     /rpt_family="L1"
     repeat_region   21009. .23092
                     /rpt_family="L1"
     repeat_region   22685. .22728
                     /rpt_family="(TG)n"
     repeat_region   23114. .23443
                     /rpt_family="L1"
     repeat_region   23523. .23592
                     /rpt_family="AT_rich"
     repeat_region   25402. .26143
                     /rpt_family="L1"
     repeat_region   26433. .26513
                     /rpt_family="CR1"
     repeat_region   27066. .27227
                     /rpt_family="MER1_type"
     repeat_region   27289. .27962
                     /rpt_family="L1"
     repeat_region   27954. .28022
                     /rpt_family="AT_rich"
     misc_feature    28774. .29234
                     /note="similar to Homo sapiens EST BF116115
                     (NID:g10985591)"
     misc_feature    28774. .29205
                     /note="similar to Homo sapiens EST BE179276
                     (NID:g8658428)"
     misc_feature    28802. .29204
                     /note="similar to Homo sapiens EST H62441 (NID:g1015273)
                     yr79d08.r1"
     misc_feature    28816. .29238
                     /note="similar to Homo sapiens EST AI079550 (NID:g3415801)
                     oz04f08.x1"
     misc_feature    28816. .29238
                     /note="similar to Homo sapiens EST AI148854 (NID:g3677323)
                     qc70a02.x1"
     misc_feature    28816. .29238

  Query Match             66.6%;  Score 675;  DB 9;  Length 104955;
  Best Local Similarity   83.0%;  Pred. No. 1.6e-155;
  Matches  793;  Conservative    0;  Mismatches  160;  Indels    2;  Gaps    2;

Qy      36 TTTCCTCTGTCATAGGGATATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACT 95
           ||   |||   | ||||| |||||||  |||   || ||||||||||||||||||||||
Db   28775 TTGTTTCTTCTACAGGGAAATGGGGGGAAATCAGACTTCCATCACAGAGTTCCTCCTACT 28834

Qy      96 GGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTA 155
           |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   28835 GGGATTTCCCATTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTA 28894

Qy     156 CGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACT 215
           | ||||||  ||| |||||||||||||| || ||||||||||||||||||||||||||||
Db   28895 CATCTTCATTCTGTTGGGGAACGGGACAATCCTGGGGCTCATCTCACTGGACTCCAGACT 28954

Qy     216 GCACGCCCCC-TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTG 274
           ||| ||||| ||||||||||||||||||||||||||||||||||||||||||  || ||
Db   28955 CCACACCCCCATGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTGTGCTTG 29014

Qy     275 CAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGC 334
           || |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   29015 CAGCACGGTGCCCCAGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGC 29074

Qy     335 GGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCT 394
           ||| ||||||  ||||  |||||||||| | ||||||    ||||||||||||||||||
Db   29075 TGGCTGCATGACCCAGATGTTTCTGTTTTTGAGTTTTGCACATACAGAATGTCTCCTCCT 29134

Qy     395 GGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCAT 454
           |||||||||||||||||||| |||||||||||||||||||| ||||||||||    ||||
Db   29135 GGTGGTGATGTCCTATGATCGGTACGTGGCCATCTGCCACCCTCTCCGATATTCTACCAT 29194

Qy     455 CATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTT 514
           |||||||||| |||||||||||  | ||  |||||||||||     |||||| | ||
Db   29195 CATGACCTGGAAAGTCTGCATCACTTTGGCATTGACTTCCTGGATTTTAGGAGTCTTATT 29254

Qy     515 ATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTA 574
```

```
Db  29255 GGCCCTTGTCCATCTAGTGTTACTGCTACCACTGTCCTTCTGTGGACCCCAGAAACTTAA 29314

Qy    575 TCACTTTTTTTTGTGAAATCTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCA 634

Db  29315 TCAC-TTTTTCTGTGAAATTATGGCTGTTCTCAAACTTGCCTGTGCGGATACCCACATTA 29373

Qy    635 ATGAGAACATGGTCTTGGCCGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTG 694

Db  29374 ATGAGGTAATGGTTTTGGCAGGGGCAGTGTCTGTGCTGGTGGGAGCCTTCTTTTCCACTG 29433

Qy    695 TAGTTTCATATATGTGCATCCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGA 754

Db  29434 TAATATCTTATGTTCATATTCTATGTGCCATTCTAAAGATCCAGTCAGGAGAGGGGTGCC 29493

Qy    755 GGAAAGCCTTCTGCACCTGCTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAG 814

Db  29494 AGAAAGCCTTCTCCATCTGCTCCTCCCACCTCTGTGTGGTTGGACTCTTTTATGGCACAG 29553

Qy    815 CCATTATCATGTATGTTGGACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCC 874

Db  29554 CCATCATCATGTATGTTGAGCCCCAGTATGAGAGCCCCAAGGAGCAGAAGAAATATCTCC 29613

Qy    875 TGCTGTTTCACAGCCTCTTTAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACT 934

Db  29614 TGCTGTTTCACAGCCTCTTCAATCCCATGCTTAATCCCCTAATTTATAGTCTTAGGAACA 29673

Qy    935 CAGAAGTGAAGAATACTTTGAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGAAA 989

Db  29674 AGGAAGTCCAAGGTACTCTAAAGAGGATGCTTGAAAAGAAGAGAACTTCATGAAA 29728


RESULT 14
AC091746
LOCUS       AC091746              204316 bp    DNA     linear   HTG 26-APR-2002
DEFINITION  Mus musculus chromosome 6 clone RP23-62E15 map 6B2, *** SEQUENCING
            IN PROGRESS ***, 30 ordered pieces.
ACCESSION   AC091746
VERSION     AC091746.2  GI:20331040
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 204316)
  AUTHORS   Young,J.M., Madan,A., Hall,J., Friedman,C., Dickhoff,R., Greene,J.,
            Lane,R.P. and Trask,B.J.
  TITLE     Evolution of the mouse olfactory receptor gene family
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 204316)
  AUTHORS   Young,J.M., Madan,A., Hall,J., Friedman,C., Dickhoff,R., Greene,J.,
            Lane,R.P. and Trask,B.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-MAY-2001) Division of Human Biology, Fred Hutchinson
            Cancer Research Center, 1100 Fairview Ave N, C3-168, P.O. Box
            19024, Seattle, WA 98109-1024, USA
REFERENCE   3  (bases 1 to 204316)
  AUTHORS   Young,J.M., Madan,A., Hall,J., Friedman,C., Dickhoff,R., Greene,J.,
            Rodrigues,S., Ketterman,M., Fahey,J., Helton,E., Sanchez,A.,
            Madan,A., Whiting,M., Ross,J.A., Lane,R.P. and Trask,B.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-APR-2002) Division of Human Biology, Fred Hutchinson
            Cancer Research Center, 1100 Fairview Ave N, C3-168, P.O. Box
            19024, Seattle, WA 98109-1024, USA
COMMENT     On Apr 26, 2002 this sequence version replaced gi:14211553.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 30 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1    11127: contig of 11127 bp in length
            *   11128    11227: gap of unknown length
            *   11228    18868: contig of 7641 bp in length
            *   18869    18968: gap of unknown length
            *   18969    40960: contig of 21992 bp in length
            *   40961    41060: gap of unknown length
            *   41061    44591: contig of 3531 bp in length
            *   44592    44691: gap of unknown length
            *   44692    46067: contig of 1376 bp in length
            *   46068    46167: gap of unknown length
            *   46168    47511: contig of 1344 bp in length
            *   47512    47611: gap of unknown length
            *   47612    51354: contig of 3743 bp in length
            *   51355    51454: gap of unknown length
            *   51455    52639: contig of 1185 bp in length
            *   52640    52739: gap of unknown length
            *   52740    54199: contig of 1460 bp in length
            *   54200    54299: gap of unknown length
            *   54300    63193: contig of 8894 bp in length
            *   63194    63293: gap of unknown length
            *   63294    70625: contig of 7332 bp in length
            *   70626    70725: gap of unknown length
            *   70726    74065: contig of 3340 bp in length
            *   74066    74165: gap of unknown length
            *   74166    77890: contig of 3725 bp in length
            *   77891    77990: gap of unknown length
            *   77991    82776: contig of 4786 bp in length
            *   82777    82876: gap of unknown length
            *   82877    90955: contig of 8079 bp in length
            *   90956    91055: gap of unknown length
            *   91056    93563: contig of 2508 bp in length
            *   93564    93663: gap of unknown length
            *   93664    98539: contig of 4876 bp in length
            *   98540    98639: gap of unknown length
            *   98640   100233: contig of 1594 bp in length
            *  100234   100333: gap of unknown length
            *  100334   108744: contig of 8411 bp in length
            *  108745   108844: gap of unknown length
            *  108845   120584: contig of 11740 bp in length
            *  120585   120684: gap of unknown length
            *  120685   122234: contig of 1550 bp in length
            *  122235   122334: gap of unknown length
            *  122335   136966: contig of 14632 bp in length
            *  136967   137066: gap of unknown length
            *  137067   146815: contig of 9749 bp in length
            *  146816   146915: gap of unknown length
            *  146916   152543: contig of 5628 bp in length
            *  152544   152643: gap of unknown length
            *  152644   164250: contig of 11607 bp in length
            *  164251   164350: gap of unknown length
            *  164351   170223: contig of 5873 bp in length
            *  170224   170323: gap of unknown length
            *  170324   189235: contig of 18912 bp in length
            *  189236   189335: gap of unknown length
            *  189336   199422: contig of 10087 bp in length
            *  199423   199522: gap of unknown length
            *  199523   201282: contig of 1760 bp in length
            *  201283   201382: gap of unknown length
            *  201383   204316: contig of 2934 bp in length.
FEATURES             Location/Qualifiers
     source          1. .204316
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="6"
                     /map="6B2"
                     /clone="RP23-62E15"
BASE COUNT    61234 a  39664 c  38680 g  61651 t   3087 others
ORIGIN

  Query Match           63.6%;  Score 645.4;  DB 2;  Length 204316;
  Best Local Similarity 80.5%;  Pred. No. 3.2e-148;
  Matches  779;  Conservative    0;  Mismatches  187;  Indels    2;  Gaps    2;
```

```
Qy      25 CATTAACTTGATTTCCTCTGTCATAGGGATATGGGGGACAATATAACATCCATCACAGAG 84
           ||| || |   ||||||||||||||||||||||||  ||||||| ||||  |||||||||
Db 117956 CATGAATTCAGTTTCCTCTGTCATAGGGATATGGGAAACAATATGACATTAATCACAGAG 118015

Qy      85 TTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTC 144
           ||| |||| ||| |||||||  |  |||| |||||  ||||||||||||||||  ||||||
Db 118016 TTCATCCTCCTGNGATTTCCTCTCAGCCCGAGGATGCAGATGCTCCTCTTTGCCCTCTTC 118075

Qy     145 TCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTG 204
           ||||| |||||  | ||||||||||||||||||  || |||||  ||||||||||||  |||
Db 118076 TCCCTCTTCTATGCCTTCACCCTGCTGGGGAATGGAACCATCGTGGGGCTCATCTGCCTG 118135

Qy     205 GACTCCAGACTGCACGC-CCCCTGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATC 263
           |||||||||||  ||| | ||| ||||||||||||| || || |||||  ||||  |||||
Db 118136 GACTCCAGACTCCACACTCCCATGTACTTCTTCCTGTCTCATCTGGCCATCGTTGACATT 118195

Qy     264 GCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCC 323
           |||||  |||||||||||  |||||||  ||||||||||||||||  |  ||||||  ||||||
Db 118196 GCCTATGCCTGCAACACAGTGCCCCAGATGCTGGTGAACCTTTTAGATCCAGTGAAGCCC 118255

Qy     324 ATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAA 383
           || |||| ||| ||  | ||||  ||||| ||||| ||||  || |||||| ||||||||
Db 118256 ATTTCCTATGCTGGTTGTATGACACAGACTTTTCTCTTTTTGACATTTGCTATCACAGAA 118315

Qy     384 TGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGA 443
           || || ||||| |||||||||||||||||  | ||| || || ||||| ||||| |||||
Db 118316 TGCCTTCTCCTAGTGGTGATGTCCTATGACCGGTATGTTGCAATCTGTCACCCACTCCGT 118375

Qy     444 TATTTGGCCATCATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACT 503
           || |  |||||||||| |||||||| |||| |||| | || ||||||||||||  | |
Db 118376 TACTCTGCCATCATGAGCTGGAGAGTTTGCAGCACCATGGCAGTGACTTCCTGGATTATT 118435

Qy     504 GGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCC 563
           || ||  |  || ||||| || || |||||||| |||| ||||||||||  | |
Db 118436 GGTGTTCTGCTGTCTTTGATCCACCTCGTGTTACTTTTACCCTTACCCTTCTGTGTGTCT 118495

Qy     564 CAGAAAATTTATCACTTTTTTTTGTGAAATCTTGGCTGTTCTCAAACTTGCCTGTGCAGA 623
           |||||| |  | ||| ||||| ||||||||    ||| | ||||||||||||||||||||
Db 118496 CAGAAAGTGAACCAC-TTTTTCTGTGAAATTACAGCTATCCTCAAACTTGCCTGTGCAGA 118554

Qy     624 TACCCACATCAATGAGAACATGGTCTTGGCCGGAGCAATTTCTGGGCTGGTGGGACCCTT 683
           || ||| |||||||||  |||||| |||| || ||| | ||||  || ||||| || ||
Db 118555 CACACACCTCAATGAGACGATGGTCCTGGCTGGGGCAGTGTCTGTCCTTGTGGGGCCATT 118614

Qy     684 GTCCACAATTGTAGTTTCATATATGTGCATCCTCTGTGCTATCCTTCAGATCCAATCAAG 743
           |||  |||||||  || || |||   || |||||  |||||||||  |||||||  |||
Db 118615 CTCCTCAATTGTGGTCTCTTATGCATGTATCCTTGGTGCTATCCTGAAGATCCAGTCAGA 118674

Qy     744 GGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTCCCACCTCTGTGTGATTGGACTCTT 803
           || | |||  |||||||||| | ||||||||  |||||||||||||||| | |||  | ||
Db 118675 AGAGGGTCAAAGGAAAGCCTTTTCCACCTGCTCCTCCCACCTCTGTGTGGTCGGATTGTT 118734

Qy     804 TTATGGCACAGCCATTATCATGTATGTTGGACCCAGATATGGGAACCCCAAGGAGCAGAA 863
           |||||| || |||||| ||||||| ||||||||||| | |||| ||| ||||||||||
Db 118735 TTATGGTACCGCCATTGTCATGTACGTTGGACCCAGACACGGGAGCCCAAAGGAGCAGAA 118794

Qy     864 GAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCCCATGCTCAATCCCCTTATCTGTAG 923
           |||||||||  |||||||||||| || ||||| || |||||||| ||| | |||| |||
Db 118795 GAAATATCTTTTGCTGTTTCACAGTCTTTTTAACCCTATGCTCAACCCCTTGATCTATAG 118854

Qy     924 TCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAGAGTGCTGGGAGTAGAAAGGGCTTT 983
           ||| ||||| || || |||||||||| ||||| |||||||| |   | | || |||||
Db 118855 TCTCAGGAATTCTGACGTGAAGAATACATTGAAAAGAGTGCTGAGGACACAGAGAGCTTT 118914

Qy     984 ATGAAAAG 991
           ||| |||
Db 118915 GTGACAAG 118922


RESULT 15
AF399598
LOCUS       AF399598                650 bp    DNA     linear   PRI 27-AUG-2001
```

```
DEFINITION  Homo sapiens clone OR2A7 olfactory receptor gene, partial cds.
ACCESSION   AF399598
VERSION     AF399598.1  GI:15293780
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 650)
  AUTHORS   Fuchs,T., Malecova,B., Linhart,C., Sharan,R., Khen,M., Herwig,R.,
            Shmulevich,D., Elkon,R., Steinfath,M., O'Brien,J.K., Radelof,U.,
            Lehrach,H., Olender,Z., Gluesman,G., Lancet,D. and Shamir,R.
  TITLE     DEFOG: A Practical Scheme for Deciphering Families of Genes
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 650)
  AUTHORS   Fuchs,T., Malecova,B., Linhart,C., Sharan,R., Khen,M., Herwig,R.,
            Shmulevich,D., Elkon,R., Steinfath,M., O'Brien,J.K., Radelof,U.,
            Lehrach,H., Olender,Z., Gluesman,G., Lancet,D. and Shamir,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUL-2001) Dept. of Molecular Genetics and the Crown
            Human Genome Center, Weizmann Institute of Science, Rehovot 76100,
            Israel
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="149.18104425"
                     /clone="OR2A7"
     mRNA            <1. .>650
                     /product="olfactory receptor"
     CDS             <1. .>650
                     /note="contains transmembrane regions 2-7"
                     /codon_start=1
                     /product="olfactory receptor"
                     /protein_id="AAK95083.1"
                     /db_xref="GI:15293781"
                     /translation="VVDIAYACNTVPRMLVNLLHPAKPISFAGRMMQTFLFSTFAVTE
                     CLLLVVMSYDLYVAICHPLRYLAIMTWRVCITLAVTSWTTGVLLSLIHLVLLLPLPFC
                     RPQKIYHFFCEILAVLKLACADTHINENMVLAGAISGLVGPLSTIVVSYMCILCAILQ
                     IQSREVQRKAFCTCFSHLCVIGLFYGTAIIMYVGPRYGNPKEQKKYLPLFHSLFNP"
BASE COUNT      131 a    187 c    136 g    196 t
ORIGIN

  Query Match           62.9%;  Score 637.4;  DB 9;  Length 650;
  Best Local Similarity  99.7%;  Pred. No. 2.3e-146;
  Matches  649;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy     252 GTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCAT 311
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCAT 60

Qy     312 CCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTT 371
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTT 120

Qy     372 GCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGC 431
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGC 180

Qy     432 CACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACT 491
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACT 240

Qy     492 TCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCC 551
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCC 300

Qy     552 TTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAAATCTTGGCTGTTCTCAAACT 611
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db     301 TTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAATCTTGGCTGTTCTCAAACT 359

Qy     612 TGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGCAATTTCTGGGCT 671
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     360 TGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGCAATTTCTGGGCT 419

Qy     672 GGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTGTGCTATCCTTCA 731
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     420 GGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTGTGCTATCCTTCA 479

Qy     732 GATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTCCCACCTCTGTGT 791
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     480 GATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTCCCACCTCTGTGT 539

Qy     792 GATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAGATATGGGAACCC 851
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     540 GATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAGATATGGGAACCC 599

Qy     852 CAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCCCAT 902
           |||||||||||||||||||| ||||||||||||||||||||||||||||
Db     600 CAAGGAGCAGAAGAAATATCTCCCGCTGTTTCACAGCCTCTTTAATCCCAT 650
```

Search completed: May 2, 2003, 05:55:22
Job time : 3242 secs

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 04:56:35 ; Search time 1671 Seconds
(without alignments)
717.034 Million cell updates/sec

Title: US-09-898-586-23
Perfect score: 1014
Sequence: 1 taaacacttctcctaaacca.....tatggcattgtgactgacaa 1014

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 746064 seqs, 590810554 residues

Total number of hits satisfying chosen parameters: 1492128

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1014	100.0	1014	9	US-09-898-586-23	Sequence 23, Appl
	2	1014	100.0	1014	10	US-09-761-288-23	Sequence 23, Appl
	3	987.4	97.4	1012	9	US-09-898-586-21	Sequence 21, Appl
	4	987.4	97.4	1012	10	US-09-761-288-21	Sequence 21, Appl
	5	983.6	97.0	1040	9	US-09-898-586-3	Sequence 3, Appli
	6	983.6	97.0	1040	9	US-09-898-586-34	Sequence 34, Appl
	7	983.6	97.0	1040	10	US-09-761-288-3	Sequence 3, Appli
	8	983.6	97.0	1040	10	US-09-761-288-34	Sequence 34, Appl
c	9	545	53.7	1957	10	US-09-864-761-13963	Sequence 13963, A
	10	538.2	53.1	2282	9	US-10-098-841-92	Sequence 92, Appl
	11	538.2	53.1	2282	10	US-09-747-835A-62	Sequence 62, Appl
	12	536.6	52.9	1788	10	US-09-747-835A-35	Sequence 35, Appl
	13	536.6	52.9	2735	10	US-09-747-835A-34	Sequence 34, Appl
c	14	535.2	52.8	933	10	US-09-864-761-30527	Sequence 30527, A
	15	511.4	50.4	1782	10	US-09-747-835A-41	Sequence 41, Appl
	16	502.4	49.5	927	10	US-09-747-835A-43	Sequence 43, Appl
	17	481.2	47.5	1713	9	US-10-024-494-1	Sequence 1, Appli
c	18	432	42.6	456	10	US-09-812-102-72	Sequence 72, Appl
	19	414	40.8	933	10	US-09-886-055-486	Sequence 486, App
	20	402.2	39.7	930	10	US-09-886-055-472	Sequence 472, App
	21	390	38.5	933	10	US-09-886-055-484	Sequence 484, App
	22	359	35.4	936	10	US-09-886-055-482	Sequence 482, App
c	23	273.2	26.9	400	9	US-09-918-995-16242	Sequence 16242, A
	24	257.2	25.4	585	10	US-09-747-835A-37	Sequence 37, Appl
	25	250.2	24.7	930	10	US-09-771-730-111	Sequence 111, App
	26	250.2	24.7	1038	10	US-09-886-055-322	Sequence 322, App
	27	249.4	24.6	921	10	US-09-771-730-135	Sequence 135, App
	28	246.4	24.3	1041	10	US-09-886-055-228	Sequence 228, App
	29	245	24.2	975	10	US-09-886-055-326	Sequence 326, App
	30	242.8	23.9	936	10	US-09-886-055-10	Sequence 10, Appl
	31	241.6	23.8	948	10	US-09-886-055-176	Sequence 176, App
c	32	240.2	23.7	1954	10	US-09-864-761-5098	Sequence 5098, Ap
	33	239.2	23.6	1031	9	US-09-974-591-5	Sequence 5, Appli
	34	239.2	23.6	1031	9	US-09-974-591-21	Sequence 21, Appl
	35	238.8	23.6	916	9	US-09-974-591-15	Sequence 15, Appl
	36	237.6	23.4	939	10	US-09-886-055-290	Sequence 290, App
	37	237.6	23.4	945	10	US-09-886-055-378	Sequence 378, App
	38	236.4	23.3	960	10	US-09-886-055-70	Sequence 70, Appl
	39	235.8	23.3	969	10	US-09-886-055-254	Sequence 254, App
	40	235.4	23.2	936	10	US-09-886-055-470	Sequence 470, App
	41	235	23.2	939	10	US-09-886-055-66	Sequence 66, Appl
	42	235	23.2	1050	9	US-09-974-591-1	Sequence 1, Appli
	43	235	23.2	1050	9	US-09-974-591-3	Sequence 3, Appli
	44	235	23.2	1050	9	US-09-974-591-7	Sequence 7, Appli
	45	234.6	23.1	942	10	US-09-886-055-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-898-586-23
; Sequence 23, Application US/09898586
; Publication No. US20030077794A1
; GENERAL INFORMATION:
; APPLICANT: Gerlach, Valerie L
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glennda
; TITLE OF INVENTION: No. US20030077794A1el Polypeptides and Nucleic Acids Encoding
; FILE REFERENCE: 15966-638CIP
; CURRENT APPLICATION NUMBER: US/09/898,586
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/177,839
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 60/176,134
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/175,989
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: 60/218,324
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/220,253
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/178,191
; PRIOR FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: 60/178,227
; PRIOR FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: 60/220,590
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 09/761,288
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-898-586-23

Query Match 100.0%; Score 1014; DB 9; Length 1014;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1014; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGG 60

Qy      61 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 120

Qy     121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 180

Qy     181 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCTGTACTTCTTCCTCT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCTGTACTTCTTCCTCT 240

Qy     241 CACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGA 300

Qy     301 ACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTGT 360

Qy     361 TTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTACG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTACG 420

Qy     421 TGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACCC 480

Qy     481 TCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTC 540

Qy     541 TACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTTGTGAAATCTTGGCT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 TACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTTGTGAAATCTTGGCT 600

Qy     601 GTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGCA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGCA 660

Qy     661 ATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTGT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 ATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTGT 720

Qy     721 GCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTCC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTCC 780

Qy     781 CACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAGA 840

Qy     841 TATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCCC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCCC 900

Qy     901 ATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAGA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAGA 960

Qy     961 GTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACAA 1014
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACAA 1014
```

RESULT 2
US-09-761-288-23
; Sequence 23, Application US/09761288
; Patent No. US20020065405A1

; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Prayaga, Sudhirdas
; APPLICANT: Taupier, Raymond J
; APPLICANT: Mishra, Vishnu
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberky
; APPLICANT: Li, Li
; TITLE OF INVENTION: No. US20020065405A1el Polypeptides and Nucleic Acids Encoding
; FILE REFERENCE: 15966-638
; CURRENT APPLICATION NUMBER: US/09/761,288
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/177,839
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 60/176,134
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/175,989
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: 60/218,324
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/220,253
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/178,191
; PRIOR FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: 60/178,227
; PRIOR FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: 60/220,590
; PRIOR FILING DATE: 2000-07-25
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-761-288-23

Query Match 100.0%; Score 1014; DB 10; Length 1014;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1014; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGG 60

Qy      61 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 120

Qy     121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 180

Qy     181 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCTGTACTTCTTCCTCT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCTGTACTTCTTCCTCT 240

Qy     241 CACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGA 300

Qy     301 ACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTGT 360

Qy     361 TTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTACG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTACG 420

Qy     421 TGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACCC 480

Qy     481 TCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTC 540
```

```
Db    481 TCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTC 540

Qy    541 TACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTTGTGAAATCTTGGCT 600

Db    541 TACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTTGTGAAATCTTGGCT 600

Qy    601 GTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGCA 660

Db    601 GTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGCA 660

Qy    661 ATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTGT 720

Db    661 ATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTGT 720

Qy    721 GCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTCC 780

Db    721 GCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTCC 780

Qy    781 CACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAGA 840

Db    781 CACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAGA 840

Qy    841 TATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCCC 900

Db    841 TATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCCC 900

Qy    901 ATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAGA 960

Db    901 ATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAGA 960

Qy    961 GTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACAA 1014

Db    961 GTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACAA 1014


RESULT 3
US-09-898-586-21
; Sequence 21, Application US/09898586
; Publication No. US20030077794A1
; GENERAL INFORMATION:
;  APPLICANT: Gerlach, Valerie L
;  APPLICANT:  MacDougall, John R
;  APPLICANT:  Smithson, Glennda
;  TITLE OF INVENTION: No. US20030077794A1el Polypeptides and Nucleic Acids Encoding Sam
;  FILE REFERENCE: 15966-638CIP
;  CURRENT APPLICATION NUMBER: US/09/898,586
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: 60/177,839
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: 60/176,134
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: 60/175,989
;  PRIOR FILING DATE: 2000-01-13
;  PRIOR APPLICATION NUMBER: 60/218,324
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/220,253
;  PRIOR FILING DATE: 2000-07-24
;  PRIOR APPLICATION NUMBER: 60/178,191
;  PRIOR FILING DATE: 2000-01-26
;  PRIOR APPLICATION NUMBER: 60/178,227
;  PRIOR FILING DATE: 2000-01-26
;  PRIOR APPLICATION NUMBER: 60/220,590
;  PRIOR FILING DATE: 2000-07-25
;  PRIOR APPLICATION NUMBER: 09/761,288
;  PRIOR FILING DATE: 2001-01-16
;  NUMBER OF SEQ ID NOS: 104
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
;   LENGTH: 1012
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-898-586-21
```

```
  Query Match             97.4%;  Score 987.4;  DB 9;  Length 1012;
  Best Local Similarity   99.7%;  Pred. No. 5.2e-312;
  Matches 1010;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    2;

Qy      2 AAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGGG 61

Db      1 AAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGAG 60

Qy     62 ACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATTC 121

Db     61 ACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATTC 120

Qy    122 AGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGGA 181

Db    121 AGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGGA 180

Qy    182 CCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTCTTCCTCT 240

Db    181 CCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTCTTCCTCT 240

Qy    241 CACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGA 300

Db    241 CACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGA 300

Qy    301 ACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTGT 360

Db    301 ACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTGT 360

Qy    361 TTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTACG 420

Db    361 TTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTACG 420

Qy    421 TGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACCC 480

Db    421 TGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACCC 480

Qy    481 TCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTC 540

Db    481 TCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTC 540

Qy    541 TACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTTGTGAAATCTTGGCT 600

Db    541 TACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAATCTTGGCT 599

Qy    601 GTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGCA 660

Db    600 GTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGCA 659

Qy    661 ATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTGT 720

Db    660 ATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTGT 719

Qy    721 GCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTCC 780

Db    720 GCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTCC 779

Qy    781 CACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAGA 840

Db    780 CACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAGA 839

Qy    841 TATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCCC 900

Db    840 TATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCCC 899

Qy    901 ATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAGA 960

Db    900 ATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAGA 959

Qy    961 GTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1013

Db    960 GTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1012
```

```
RESULT 4
US-09-761-288-21
; Sequence 21, Application US/09761288
; Patent No. US20020065405A1
; GENERAL INFORMATION:
;  APPLICANT: Padigaru, Muralidhara
;  APPLICANT:  Prayaga, Sudhirdas
;  APPLICANT:  Taupier, Raymond J
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Spytek, Kimberky
;  APPLICANT:  Li, Li
;  TITLE OF INVENTION: No. US20020065405A1el Polypeptides and Nucleic Acids Encoding Sam
;  FILE REFERENCE: 15966-638
;  CURRENT APPLICATION NUMBER: US/09/761,288
;  CURRENT FILING DATE:  2001-01-16
;  PRIOR APPLICATION NUMBER: 60/177,839
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: 60/176,134
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: 60/175,989
;  PRIOR FILING DATE: 2000-01-13
;  PRIOR APPLICATION NUMBER: 60/218,324
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/220,253
;  PRIOR FILING DATE: 2000-07-24
;  PRIOR APPLICATION NUMBER: 60/178,191
;  PRIOR FILING DATE: 2000-01-26
;  PRIOR APPLICATION NUMBER: 60/178,227
;  PRIOR FILING DATE: 2000-01-26
;  PRIOR APPLICATION NUMBER: 60/220,590
;  PRIOR FILING DATE: 2000-07-25
;  NUMBER OF SEQ ID NOS: 95
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
;   LENGTH: 1012
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-761-288-21

  Query Match             97.4%;  Score 987.4;  DB 10;  Length 1012;
  Best Local Similarity   99.7%;  Pred. No. 5.2e-312;
  Matches 1010;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    2;

Qy       2 AAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGGG 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1 AAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGAG 60

Qy      62 ACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATTC 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATTC 120

Qy     122 AGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGGA 181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGGA 180

Qy     182 CCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTCTTCCTCT 240
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     181 CCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTCTTCCTCT 240

Qy     241 CACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGA 300

Qy     301 ACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTGT 360

Qy     361 TTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTACG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTACG 420

Qy     421 TGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACCC 480

Qy     481 TCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTC 540

Qy     541 TACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAAATCTTGGCT 600
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     541 TACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAATCTTGGCT 599

Qy     601 GTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGCA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 GTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGCA 659

Qy     661 ATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTGT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     660 ATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTGT 719

Qy     721 GCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTCC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     720 GCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTCC 779

Qy     781 CACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     780 CACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAGA 839

Qy     841 TATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCCC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     840 TATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCCC 899

Qy     901 ATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAGA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     900 ATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAGA 959

Qy     961 GTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1013
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     960 GTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1012


RESULT 5
US-09-898-586-3
; Sequence 3, Application US/09898586
; Publication No. US20030077794A1
; GENERAL INFORMATION:
;  APPLICANT: Gerlach, Valerie L
;  APPLICANT:  MacDougall, John R
;  APPLICANT:  Smithson, Glennda
;  TITLE OF INVENTION: No. US20030077794A1el Polypeptides and Nucleic Acids Encoding
;  FILE REFERENCE: 15966-638CIP
;  CURRENT APPLICATION NUMBER: US/09/898,586
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: 60/177,839
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: 60/176,134
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: 60/175,989
;  PRIOR FILING DATE: 2000-01-13
;  PRIOR APPLICATION NUMBER: 60/218,324
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/220,253
;  PRIOR FILING DATE: 2000-07-24
;  PRIOR APPLICATION NUMBER: 60/178,191
;  PRIOR FILING DATE: 2000-01-26
;  PRIOR APPLICATION NUMBER: 60/178,227
;  PRIOR FILING DATE: 2000-01-26
;  PRIOR APPLICATION NUMBER: 60/220,590
;  PRIOR FILING DATE: 2000-07-25
;  PRIOR APPLICATION NUMBER: 09/761,288
;  PRIOR FILING DATE: 2001-01-16
;  NUMBER OF SEQ ID NOS: 104
;  SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 3
;   LENGTH: 1040
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-898-586-3

  Query Match             97.0%;  Score 983.6;  DB 9;  Length 1040;
  Best Local Similarity   99.4%;  Pred. No. 9.2e-311;
  Matches 1008;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy       1 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      28 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGA 87

Qy      61 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 120
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      88 GACAATATAACATCCATCAGAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 147

Qy     121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     148 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 207

Qy     181 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTCTTCCTC 239
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     208 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTCTTCCTC 267

Qy     240 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTG 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     268 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTG 327

Qy     300 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     328 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 387

Qy     360 TTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     388 TTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 447

Qy     420 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     448 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 507

Qy     480 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     508 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 567

Qy     540 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAAATCTTGGC 599
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     568 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAATCTTGGC 626

Qy     600 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     627 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 686

Qy     660 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     687 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 746

Qy     720 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTC 779
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     747 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCCGCACCTGCTTCTC 806

Qy     780 CCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 839
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db     807 CCACCTCTGTGTGATTGGACTCGTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 866

Qy     840 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     867 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 926

Qy     900 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     927 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 986

Qy     960 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1013
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     987 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1040


RESULT 6
US-09-898-586-34
; Sequence 34, Application US/09898586
; Publication No. US20030077794A1
; GENERAL INFORMATION:
;  APPLICANT: Gerlach, Valerie L
;  APPLICANT:  MacDougall, John R
;  APPLICANT:  Smithson, Glennda
;  TITLE OF INVENTION: No. US20030077794Alel Polypeptides and Nucleic Acids Encoding
;  FILE REFERENCE: 15966-638CIP
;  CURRENT APPLICATION NUMBER: US/09/898,586
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: 60/177,839
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: 60/176,134
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: 60/175,989
;  PRIOR FILING DATE: 2000-01-13
;  PRIOR APPLICATION NUMBER: 60/218,324
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/220,253
;  PRIOR FILING DATE: 2000-07-24
;  PRIOR APPLICATION NUMBER: 60/178,191
;  PRIOR FILING DATE: 2000-01-26
;  PRIOR APPLICATION NUMBER: 60/178,227
;  PRIOR FILING DATE: 2000-01-26
;  PRIOR APPLICATION NUMBER: 60/220,590
;  PRIOR FILING DATE: 2000-07-25
;  PRIOR APPLICATION NUMBER: 09/761,288
;  PRIOR FILING DATE: 2001-01-16
;  NUMBER OF SEQ ID NOS: 104
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
;   LENGTH: 1040
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-898-586-34

  Query Match             97.0%;  Score 983.6;  DB 9;  Length 1040;
  Best Local Similarity   99.4%;  Pred. No. 9.2e-311;
  Matches 1008;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy       1 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      28 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGA 87

Qy      61 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 120
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      88 GACAATATAACATCCATCAGAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 147

Qy     121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     148 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 207

Qy     181 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTCTTCCTC 239
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     208 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTCTTCCTC 267

Qy     240 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTG 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     268 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTG 327

Qy     300 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     328 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 387

Qy     360 TTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 419
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 TTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 447

Qy        420 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 507

Qy        480 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 567

Qy        540 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAAATCTTGGC 599
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        568 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAATCTTGGC 626

Qy        600 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 686

Qy        660 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 746

Qy        720 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTC 779
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        747 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCCGCACCTGCTTCTC 806

Qy        780 CCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 839
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        807 CCACCTCTGTGTGATTGGACTCGTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 866

Qy        840 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 899
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 926

Qy        900 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 959
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 986

Qy        960 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1013
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1040
```

RESULT 7
US-09-761-288-3
; Sequence 3, Application US/09761288
; Patent No. US20020065405A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Prayaga, Sudhirdas
; APPLICANT: Taupier, Raymond J
; APPLICANT: Mishra, Vishnu
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberky
; APPLICANT: Li, Li
; TITLE OF INVENTION: No. US20020065405Alel Polypeptides and Nucleic Acids Encoding Sam
; FILE REFERENCE: 15966-638
; CURRENT APPLICATION NUMBER: US/09/761,288
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/177,839
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 60/176,134
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/175,989
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: 60/218,324
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/220,253
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/178,191
; PRIOR FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: 60/178,227
; PRIOR FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: 60/220,590
; PRIOR FILING DATE: 2000-07-25
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1040
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-761-288-3

```
  Query Match             97.0%;  Score 983.6;  DB 10;  Length 1040;
  Best Local Similarity   99.4%;  Pred. No. 9.2e-311;
  Matches 1008;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy          1 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGA 87

Qy         61 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 120
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         88 GACAATATAACATCCATCAGAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 147

Qy        121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 207

Qy        181 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTCTTCCTC 239
              ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db        208 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTCTTCCTC 267

Qy        240 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTG 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTG 327

Qy        300 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 387

Qy        360 TTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 TTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 447

Qy        420 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 507

Qy        480 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 567

Qy        540 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAAATCTTGGC 599
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        568 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAATCTTGGC 626

Qy        600 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 686

Qy        660 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 746

Qy        720 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTC 779
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        747 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCCGCACCTGCTTCTC 806

Qy        780 CCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 839
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        807 CCACCTCTGTGTGATTGGACTCGTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 866

Qy        840 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 899
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    867 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 926

Qy    900 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    927 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 986

Qy    960 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1013
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    987 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1040


RESULT 8
US-09-761-288-34
; Sequence 34, Application US/09761288
; Patent No. US20020065405A1
; GENERAL INFORMATION:
;  APPLICANT: Padigaru, Muralidhara
;  APPLICANT:  Prayaga, Sudhirdas
;  APPLICANT:  Taupier, Raymond J
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Spytek, Kimberky
;  APPLICANT:  Li, Li
;  TITLE OF INVENTION: No. US20020065405A1el Polypeptides and Nucleic Acids Encoding Sam
;  FILE REFERENCE: 15966-638
;  CURRENT APPLICATION NUMBER: US/09/761,288
;  CURRENT FILING DATE:  2001-01-16
;  PRIOR APPLICATION NUMBER: 60/177,839
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: 60/176,134
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: 60/175,989
;  PRIOR FILING DATE: 2000-01-13
;  PRIOR APPLICATION NUMBER: 60/218,324
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/220,253
;  PRIOR FILING DATE: 2000-07-24
;  PRIOR APPLICATION NUMBER: 60/178,191
;  PRIOR FILING DATE: 2000-01-26
;  PRIOR APPLICATION NUMBER: 60/178,227
;  PRIOR FILING DATE: 2000-01-26
;  PRIOR APPLICATION NUMBER: 60/220,590
;  PRIOR FILING DATE: 2000-07-25
;  NUMBER OF SEQ ID NOS: 95
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
;   LENGTH: 1040
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-761-288-34

  Query Match             97.0%;  Score 983.6;  DB 10;  Length 1040;
  Best Local Similarity   99.4%;  Pred. No. 9.2e-311;
  Matches 1008;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy      1 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     28 TAAACACTTCTCCTAAACCATGAGCATTAACTTGATTTCCTCTGTCATAGGGATATGGGA 87

Qy     61 GACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 120
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     88 GACAATATAACATCCATCAGAGAGTTCCTCCTACTGGGATTTCCCGTTGGCCCAAGGATT 147

Qy    121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    148 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTGGGGAACGGG 207

Qy    181 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTCTTCCTC 239
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    208 ACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTCTTCCTC 267

Qy    240 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTG 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    268 TCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTG 327

Qy    300 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    328 AACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACCTTTCTG 387

Qy    360 TTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    388 TTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTATGATCTGTAC 447

Qy    420 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    448 GTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCACC 507

Qy    480 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    508 CTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTT 567

Qy    540 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAAATCTTGGC 599
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    568 CTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAAATCTTGGC 626

Qy    600 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    627 TGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCCGGAGC 686

Qy    660 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    687 AATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTGCATCCTCTG 746

Qy    720 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTGCTTCTC 779
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    747 TGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCCGCACCTGCTTCTC 806

Qy    780 CCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 839
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    807 CCACCTCTGTGTGATTGGACTCGTTTATGGCACAGCCATTATCATGTATGTTGGACCCAG 866

Qy    840 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    867 ATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCC 926

Qy    900 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    927 CATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTGAAGAG 986

Qy    960 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1013
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    987 AGTGCTGGGAGTAGAAAGGGCTTTATGAAAAGGATTATGGCATTGTGACTGACA 1040


RESULT 9
US-09-864-761-13963/c
; Sequence 13963, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL F
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
```

```
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 13963
;   LENGTH: 1957
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC004889.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.93
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
US-09-864-761-13963

  Query Match             53.7%;  Score 545;  DB 10;  Length 1957;
  Best Local Similarity   74.6%;  Pred. No. 3.1e-167;
  Matches  711;  Conservative    0;  Mismatches  240;  Indels    2;  Gaps    2;

Qy      36 TTTCCTCTGTCATAGGGATATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACT 95
           ||| || | ||| ||||| |||||||| |||   |||    ||||||||||||||||||
Db    1073 TTTTCTTTTTCACAGGGAAATGGGGGAAAATCAGACAATGGTCACAGAGTTCCTCCTACT 1014

Qy      96 GGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTA 155
           |||||||| | | ||||||||||||||||||||||||||||||||||||||||||||||
Db    1013 GGGATTTCTCCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTA 954

Qy     156 CGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACT 215
           ||||||||||||||||||||||| |||| ||||||||||||||||||||||||||||||
Db     953 TATCTTCACCCTGCTGGGGAACGGGGCCATCCTGGGGCTCATCTCACTGGACTCCAGACT 894

Qy     216 GCACGCCCCC-TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTG 274
           ||| ||||| |||||||||||||||||||||||| |||||||||||||||||| || |
Db     893 CCACACCCCCATGTACTTCTTCCTCTCACACCTGGCTGTCGTCGACATCGCCTACACCCG 834

Qy     275 CAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGC 334
           ||||||||||||||| |||||||| |||||||||||||||||||||||||||||||||||
Db     833 CAACACGGTGCCCCAGATGCTGGCGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGC 774

Qy     335 GGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCT 394
           ||  |||||| |||||||||||| | ||  | |||||    ||  ||||||||||| ||
Db     773 TGGTTGCATGACGCAGACCTTTCTCTGTTTGAGTTTTGGACACAGCGAATGTCTCCTGCT 714

Qy     395 GGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCAT 454
           |||| |||||||||| |||| ||||||||||||||||||||| |||||||| |  | |||
Db     713 GGTGCTGATGTCCTACGATCGTTACGTGGCCATCTGCCACCCTCTCCGATACTCCGTCAT 654

Qy     455 CATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTT 514
           ||||||||||||||||||||||||||||| || || |||||||||||   |||   ||| |
Db     653 CATGACCTGGAGAGTCTGCATCACCCTGGCCGTCACTTCCTGGACGTGTGGCTCCCTCCT 594

Qy     515 ATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTA 574
           |  ||   ||| | ||  |  | |||    | ||||||| | |||| ||  ||||  |
Db     593 GGCTCTGGCCCATGTGGTTCTCATCCTAAGACTGCCCTTCTCTGGGCCTCATGAAATCAA 534

Qy     575 TCACTTTTTTTTGTGAAATCTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCA 634
           ||| || || |||||||||| || |||| ||||  || ||||||||| || |||    |||
Db     533 CCAC-TTCTTCTGTGAAATCCTGTCTGTCCTCAGGCTGGCCTGTGCTGACACCTGGCTCA 475

Qy     635 ATGAGAACATGGTCTTGGCCGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTG 694
           |  ||    |  |||| || |      | |      |||||||| ||       |   |
Db     474 ACCAGGTGGTCATCTTTGCAGCCTGCGTGTTCTTCCTGGTGGGGCCACCCAGCCTGGTGC 415

Qy     695 TAGTTTCATATATGTGCATCCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGA 754
           | || || ||   |  ||||||    || |||||  ||||||  ||  |||| |  |  |
Db     414 TTGTCTCCTACTCGCACATCCTGGCGGCCATCCTGAGGATCCAGTCTGGGGAGGGCCGCA 355

Qy     755 GGAAAGCCTTCTGCACCTGCTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAG 814
           | || |||||||  |||||||| |||||||||||| ||| | |||||||| | |||||  |
Db     354 GAAAGGCCTTCTCCACCTGCTCCTCCCACCTCTGCGTGGTGGGACTCTTCTTTGGCAGTG 295

Qy     815 CCATTATCATGTATGTTGGACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCC 874
           |||| ||||||||  | |  |||| |   |  | ||  |||||||| | ||     |
Db     294 CCATCATCATGTACATGGCCCCCAAGTCCCGCCATCCTGAGGAGCAGCAAAAGGTCTTTT 235

Qy     875 TGCTGTTTCACAGCCTCTTTAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACT 934
           | || ||| ||||  | || || || |  || || ||||| || |  || || ||||||
Db     234 TTCTATTTTACAGTTTTTTCAACCCAACACTTAACCCCCTGATTTACAGCCTGAGGAACG 175

Qy     935 CAGAAGTGAAGAATACTTTGAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGA 987
           ||| || |||  | |  ||| |||||  |||||    |||||   ||  | |
Db     174 GAGAGGTCAAGGGTGCCCTGAGGAGAGCACTGGGCAAGGAAAGTCATTCCTAA 122


RESULT 10
US-10-098-841-92
; Sequence 92, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Xu, Chongjun
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Qian, Xiaohong B.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. US20020197679Alel Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2
;  CURRENT APPLICATION NUMBER: US/10/098,841
;  CURRENT FILING DATE:  2002-03-13
;  PRIOR APPLICATION NUMBER: 09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 331
```

```
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 92
;   LENGTH: 2282
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1257)..(2189)
US-10-098-841-92

  Query Match            53.1%;  Score 538.2;  DB 9;  Length 2282;
  Best Local Similarity  73.6%;  Pred. No. 5.6e-165;
  Matches  712;  Conservative    0;  Mismatches  253;  Indels    2;  Gaps    2;

Qy      31 CTTGATTTCCTCTGTCATAGGGATATGGGGGACAATATAACATCCATCACAGAGTTCCTC 90
           |||  |||  | | |  ||| || |||| | | |||   |||    ||||||||||||||
Db    1233 CTTTGTTTGTTTTTTAGTAGTGAAATGGTGAAAAATCAGACAATGGTCACAGAGTTCCTC 1292

Qy      91 CTACTGGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 150
           |||||||||||||  | | ||||||||||||||||||||||||||||||||||||||||
Db    1293 CTACTGGGATTTCTCCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 1352

Qy     151 TTCTACGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCC 210
           |||||  |||||||||||||||||||  |||||||| ||||||||||||||||||||||
Db    1353 TTCTATGTCTTCACCCTGCTGGGGAATGGGACCATCCTGGGGCTCATCTCACTGGACTCC 1412

Qy     211 AGACTGCACGCCCCC-TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTAC 269
           |||||  |||  |||||  |||||||||||||||||||||||||  ||||||  ||||||||||
Db    1413 AGACTCCACACCCCCATGTACTTCTTCCTCTCACACCTGGCCGTCGTCAACATCGCCTAT 1472

Qy     270 GCCTGCAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCC 329
           ||||||||||| |||||||  |||||||||||||||||||||||||||||||||||||||||
Db    1473 GCCTGCAACACAGTGCCCCAGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCC 1532

Qy     330 TTTGCGGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTC 389
           |||||  |||  ||||||  ||||||||||| ||||  | ||||||     || |||||  |||
Db    1533 TTTGCTGGCTGCATGACACAGACCTTTCTCTTTTTGAGTTTTGCACATACTGAATGCCTC 1592

Qy     390 CTCCTGGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTG 449
           ||  |||||  ||||||||||  |||| ||||||||||||||||||||  ||||||||||
Db    1593 CTGTTGGTGCTGATGTCCTACGATCGGTACGTGGCCATCTGCCACCCTCTCCGATATTTC 1652

Qy     450 GCCATCATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTC 509
             ||||||||||||| ||||||||||| || ||  | ||||||||||   |||   |
Db    1653 ATCATCATGACCTGGAAAGTCTGCATCACTCTGGCCATCACTTCCTGGACATGTGGCTCC 1712

Qy     510 CTTTTATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAA 569
           ||  |  |  || | ||| |     |  | |||    | |||||  ||| |||| |   ||
Db    1713 CTCCTGGCTATGGTCCATGTGAGCCTCATCCTAAGACTGCCCTTTTGTGGGCCTCGTGAA 1772

Qy     570 ATTTATCACTTTTTTTTTGTGAAATCTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCA 629
           ||  | ||| || || |||||||||| || |||| ||||  || ||||||||  ||||||
Db    1773 ATCAACCAC-TTCTTCTGTGAAATCCTGTCTGTCCTCAGGCTGGCCTGTGCTGATACCTG 1831

Qy     630 CATCAATGAGAACATGGTCTTGGCCGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCAC 689
             ||||  ||     |  |||| || |     || |     ||||||||||  | | |
Db    1832 GCTCAACCAGGTGGTCATCTTTGCAGCCTGCATGTTCATCCTGGTGGGACCACTCTGCCT 1891

Qy     690 AATTGTAGTTTCATATATGTGCATCCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGT 749
             |  | || || ||       ||||||    || |||||   ||||||  ||  |||| |
Db    1892 GGTGCTGGTCTCCTACTCACACATCCTGGCGGCCATCCTGAGGATCCAGTCTGGGGAGGG 1951

Qy     750 TCAGAGGAAAGCCTTCTGCACCTGCTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGG 809
             |  || || ||||||| |||||||| ||||||||||| ||  | ||||||| | |||
Db    1952 CCGCAGAAAGGCCTTCTCCACCTGCTCCTCCCACCTCTGCGTAGTGGGACTCTTCTTTGG 2011

Qy     810 CACAGCCATTATCATGTATGTTGGACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATA 869
           ||  |||||  |||||||  | |  || |  |   | | ||  |||||||| ||||
Db    2012 CAGCGCCATCGTCATGTACATGGCCCCTAAGTCCCGCCATCCTGAGGAGCAGCAGAAGGT 2071

Qy     870 TCTCCTGCTGTTTCACAGCCTCTTTAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAG 929
             ||  | || ||| ||||    || || || ||||| || ||||| || |  |  || ||
Db    2072 CCTTTTTCTATTTTACAGTTCTTTCAACCCGATGCTAAACCCCCTGATTTACAACCTGAG 2131

Qy     930 GAACTCAGAAGTGAAGAATACTTTGAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGAAA 989
           |||   ||| || |||  | |  ||| |||||  ||| |    |||||   ||  | | |
Db    2132 GAATGTAGAGGTCAAGGGTGCCCTGAGGAGAGCACTGTGCAAGGAAAGTCATTCCTAAGA 2191

Qy     990 AGGATTA 996
           |  | |
Db    2192 GGTGTGA 2198


RESULT 11
US-09-747-835A-62
; Sequence 62, Application US/09747835A
; Patent No. US20020146692A1
; GENERAL INFORMATION:
;  APPLICANT: Yamazaki, Victoria
;  APPLICANT:  Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Drmanac, Radoje T
;  TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-
;  TITLE OF INVENTION:  LIKE) POLYPEPTIDES AND POLYNUCLEOTIDES
;  FILE REFERENCE: HYS-37CIP
;  CURRENT APPLICATION NUMBER: US/09/747,835A
;  CURRENT FILING DATE:  2002-03-08
;  PRIOR APPLICATION NUMBER: US 09/729,739
;  PRIOR FILING DATE: 2000-12-04
;  PRIOR APPLICATION NUMBER: US 09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US 09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US 09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 62
;   LENGTH: 2282
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1257)..(2189)
US-09-747-835A-62

  Query Match            53.1%;  Score 538.2;  DB 10;  Length 2282;
  Best Local Similarity  73.6%;  Pred. No. 5.6e-165;
  Matches  712;  Conservative    0;  Mismatches  253;  Indels    2;  Gaps    2;

Qy      31 CTTGATTTCCTCTGTCATAGGGATATGGGGGACAATATAACATCCATCACAGAGTTCCTC 90
           |||  |||  | | |  ||| || |||| | | |||   |||    ||||||||||||||
Db    1233 CTTTGTTTGTTTTTTAGTAGTGAAATGGTGAAAAATCAGACAATGGTCACAGAGTTCCTC 1292

Qy      91 CTACTGGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 150
           |||||||||||||  | | ||||||||||||||||||||||||||||||||||||||||
Db    1293 CTACTGGGATTTCTCCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 1352

Qy     151 TTCTACGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCC 210
           |||||  |||||||||||||||||||  |||||||| ||||||||||||||||||||||
Db    1353 TTCTATGTCTTCACCCTGCTGGGGAATGGGACCATCCTGGGGCTCATCTCACTGGACTCC 1412

Qy     211 AGACTGCACGCCCCC-TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTAC 269
           |||||  |||  |||||  |||||||||||||||||||||||||  ||||||  ||||||||||
Db    1413 AGACTCCACACCCCCATGTACTTCTTCCTCTCACACCTGGCCGTCGTCAACATCGCCTAT 1472

Qy     270 GCCTGCAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCC 329
```

```
Db   1473 GCCTGCAACACAGTGCCCCAGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCC 1532

Qy    330 TTTGCGGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTC 389

Db   1533 TTTGCTGGCTGCATGACACAGACCTTTCTCTTTTTGAGTTTTGCACATACTGAATGCCTC 1592

Qy    390 CTCCTGGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTG 449

Db   1593 CTGTTGGTGCTGATGTCCTACGATCGGTACGTGGCCATCTGCCACCCTCTCCGATATTTC 1652

Qy    450 GCCATCATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTC 509

Db   1653 ATCATCATGACCTGGAAAGTCTGCATCACTCTGGCCATCACTTCCTGGACATGTGGCTCC 1712

Qy    510 CTTTTATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAA 569

Db   1713 CTCCTGGCTATGGTCCATGTGAGCCTCATCCTAAGACTGCCCTTTTGTGGGCCTCGTGAA 1772

Qy    570 ATTTATCACTTTTTTTTGTGAAATCTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCA 629

Db   1773 ATCAACCAC-TTCTTCTGTGAAATCCTGTCTGTCCTCAGGCTGGCCTGTGCTGATACCTG 1831

Qy    630 CATCAATGAGAACATGGTCTTGGCCGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCAC 689

Db   1832 GCTCAACCAGGTGGTCATCTTTGCAGCCTGCATGTTCATCCTGGTGGGACCACTCTGCCT 1891

Qy    690 AATTGTAGTTTCATATATGTGCATCCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGT 749

Db   1892 GGTGCTGGTCTCCTACTCACACATCCTGGCGGCCATCCTGAGGATCCAGTCTGGGGAGGG 1951

Qy    750 TCAGAGGAAAGCCTTCTGCACCTGCTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGG 809

Db   1952 CCGCAGAAAGGCCTTCTCCACCTGCTCCTCCCACCTCTGCGTAGTGGGACTCTTCTTTGG 2011

Qy    810 CACAGCCATTATCATGTATGTTGGACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATA 869

Db   2012 CAGCGCCATCGTCATGTACATGGCCCCTAAGTCCCGCCATCCTGAGGAGCAGCAGAAGGT 2071

Qy    870 TCTCCTGCTGTTTCACAGCCTCTTTAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAG 929

Db   2072 CCTTTTTCTATTTTACAGTTCTTTCAACCCGATGCTAAACCCCCTGATTTACAACCTGAG 2131

Qy    930 GAACTCAGAAGTGAAGAATACTTTGAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGAAA 989

Db   2132 GAATGTAGAGGTCAAGGGTGCCCTGAGGAGAGCACTGTGCAAGGAAAGTCATTCCTAAGA 2191

Qy    990 AGGATTA 996

Db   2192 GGTGTGA 2198
```

RESULT 12
US-09-747-835A-35
; Sequence 35, Application US/09747835A
; Patent No. US20020146692A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
; TITLE OF INVENTION: LIKE) POLYPEPTIDES AND POLYNUCLEOTIDES
; FILE REFERENCE: HYS-37CIP
; CURRENT APPLICATION NUMBER: US/09/747,835A
; CURRENT FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 35
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (833)..(1417)
US-09-747-835A-35

Query Match 52.9%; Score 536.6; DB 10; Length 1788;
Best Local Similarity 73.5%; Pred. No. 1.6e-164;
Matches 711; Conservative 0; Mismatches 254; Indels 2; Gaps 2;

```
Qy     31 CTTGATTTCCTCTGTCATAGGGATATGGGGGACAATATAACATCCATCACAGAGTTCCTC 90

Db    461 CTTTGTTTGTTTTTTAGTAGTGAAATGGTGAAAAATCAGACAATGGTCACAGAGTTCCTC 520

Qy     91 CTACTGGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 150

Db    521 CTACTGGGATTTCTCCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 580

Qy    151 TTCTACGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCC 210

Db    581 TTCTATGTCTTCACCCTGCTGGGGAATGGGACCATCCTGGGGCTCATCTCACTGGACTCC 640

Qy    211 AGACTGCACGCCCCC-TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTAC 269

Db    641 AGACTCCACACCCCCATGTACTTCTTCCTCTCACACCTGGCCGTCGTCAACATCGCCTAT 700

Qy    270 GCCTGCAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCC 329

Db    701 GCCTGCAACACAGTGCCCCAGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCC 760

Qy    330 TTTGCGGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTC 389

Db    761 TTTGCTGGCTGCATGACATAGACCTTTCTCTTTTTGAGTTTTGCACATACTGAATGCCTC 820

Qy    390 CTCCTGGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTG 449

Db    821 CTGTTGGTGCTGATGTCCTACGATCGGTACGTGGCCATCTGCCACCCTCTCCGATATTTC 880

Qy    450 GCCATCATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTC 509

Db    881 ATCATCATGACCTGGAAAGTCTGCATCACTCTGGCCATCACTTCCTGGACATGTGGCTCC 940

Qy    510 CTTTTATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAA 569

Db    941 CTCCTGGCTATGGTCCATGTGAGCCTCATCCTAAGACTGCCCTTTTGTGGGCCTCGTGAA 1000

Qy    570 ATTTATCACTTTTTTTTGTGAAATCTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCA 629

Db   1001 ATCAACCAC-TTCTTCTGTGAAATCCTGTCTGTCCTCAGGCTGGCCTGTGCTGATACCTG 1059

Qy    630 CATCAATGAGAACATGGTCTTGGCCGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCAC 689

Db   1060 GCTCAACCAGGTGGTCATCTTTGCAGCCTGCATGTTCATCCTGGTGGGACCACTCTGCCT 1119

Qy    690 AATTGTAGTTTCATATATGTGCATCCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGT 749

Db   1120 GGTGCTGGTCTCCTACTCACACATCCTGGCGGCCATCCTGAGGATCCAGTCTGGGGAGGG 1179
```

```
Qy     750 TCAGAGGAAAGCCTTCTGCACCTGCTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGG 809
Db    1180 CCGCAGAAAGGCCTTCTCCACCTGCTCCTCCCACCTCTGCGTAGTGGGACTCTTCTTTGG 1239

Qy     810 CACAGCCATTATCATGTATGTTGGACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATA 869
Db    1240 CAGCGCCATCGTCATGTACATGGCCCCTAAGTCCCGCCATCCTGAGGAGCAGCAGAAGGT 1299

Qy     870 TCTCCTGCTGTTTCACAGCCTCTTTAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAG 929
Db    1300 CCTTTTTCTATTTTACAGTTCTTTCAACCCGATGCTAAACCCCCTGATTTACAACCTGAG 1359

Qy     930 GAACTCAGAAGTGAAGAATACTTTGAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGAAA 989
Db    1360 GAATGTAGAGGTCAAGGGTGCCCTGAGGAGAGCACTGTGCAAGGAAAGTCATTCCTAAGA 1419

Qy     990 AGGATTA 996
Db    1420 GGTGTGA 1426
```

RESULT 13
US-09-747-835A-34
; Sequence 34, Application US/09747835A
; Patent No. US20020146692A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
; TITLE OF INVENTION: LIKE) POLYPEPTIDES AND POLYNUCLEOTIDES
; FILE REFERENCE: HYS-37CIP
; CURRENT APPLICATION NUMBER: US/09/747,835A
; CURRENT FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 2735
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-747-835A-34

Query Match 52.9%; Score 536.6; DB 10; Length 2735;
Best Local Similarity 73.5%; Pred. No. 2.1e-164;
Matches 711; Conservative 0; Mismatches 254; Indels 2; Gaps 2;

```
Qy      31 CTTGATTTCCTCTGTCATAGGGATATGGGGGACAATATAACATCCATCACAGAGTTCCTC 90
Db    1311 CTTTGTTTGTTTTTTAGTAGTGAAATGGTGAAAAATCAGACAATGGTCACAGAGTTCCTC 1370

Qy      91 CTACTGGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 150
Db    1371 CTACTGGGATTTCTCCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 1430

Qy     151 TTCTACGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCC 210
Db    1431 TTCTATGTCTTCACCCTGCTGGGGAATGGGACCATCCTGGGGCTCATCTCACTGGACTCC 1490

Qy     211 AGACTGCACGCCCCC-TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTAC 269
Db    1491 AGACTCCACACCCCCATGTACTTCTTCCTCTCACACCTGGCCGTCGTCAACATCGCCTAT 1550

Qy     270 GCCTGCAACACGGTGCCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCC 329
Db    1551 GCCTGCAACACAGTGCCCCAGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCC 1610

Qy     330 TTTGCGGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTC 389
Db    1611 TTTGCTGGCTGCATGACATAGACCTTTCTCTTTTTGAGTTTTGCACATACTGAATGCCTC 1670

Qy     390 CTCCTGGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTG 449
Db    1671 CTGTTGGTGCTGATGTCCTACGATCGGTACGTGGCCATCTGCCACCCTCTCCGATATTTC 1730

Qy     450 GCCATCATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTC 509
Db    1731 ATCATCATGACCTGGAAAGTCTGCATCACTCTGGCCATCACTTCCTGGACATGTGGCTCC 1790

Qy     510 CTTTTATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAA 569
Db    1791 CTCCTGGCTATGGTCCATGTGAGCCTCATCCTAAGACTGCCCTTTTGTGGGCCTCGTGAA 1850

Qy     570 ATTTATCACTTTTTTTTGTGAAATCTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCA 629
Db    1851 ATCAACCAC-TTCTTCTGTGAAATCCTGTCTGTCCTCAGGCTGGCCTGTGCTGATACCTG 1909

Qy     630 CATCAATGAGAACATGGTCTTGGCCGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCAC 689
Db    1910 GCTCAACCAGGTGGTCATCTTTGCAGCCTGCATGTTCATCCTGGTGGGACCACTCTGCCT 1969

Qy     690 AATTGTAGTTTCATATATGTGCATCCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGT 749
Db    1970 GGTGCTGGTCTCCTACTCACACATCCTGGCGGCCATCCTGAGGATCCAGTCTGGGGAGGG 2029

Qy     750 TCAGAGGAAAGCCTTCTGCACCTGCTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGG 809
Db    2030 CCGCAGAAAGGCCTTCTCCACCTGCTCCTCCCACCTCTGCGTAGTGGGACTCTTCTTTGG 2089

Qy     810 CACAGCCATTATCATGTATGTTGGACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATA 869
Db    2090 CAGCGCCATCGTCATGTACATGGCCCCTAAGTCCCGCCATCCTGAGGAGCAGCAGAAGGT 2149

Qy     870 TCTCCTGCTGTTTCACAGCCTCTTTAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAG 929
Db    2150 CCTTTTTCTATTTTACAGTTCTTTCAACCCGATGCTAAACCCCCTGATTTACAACCTGAG 2209

Qy     930 GAACTCAGAAGTGAAGAATACTTTGAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGAAA 989
Db    2210 GAATGTAGAGGTCAAGGGTGCCCTGAGGAGAGCACTGTGCAAGGAAAGTCATTCCTAAGA 2269

Qy     990 AGGATTA 996
Db    2270 GGTGTGA 2276
```

RESULT 14
US-09-864-761-30527/c
; Sequence 30527, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL F
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761

```
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30527
;   LENGTH: 933
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC004889.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.93
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
;   OTHER INFORMATION: SWISSPROT HIT: P34984, EVALUE 1.00e-86
;   OTHER INFORMATION: EST_HUMAN HIT: BF116115.1, EVALUE 0.00e+00
;   OTHER INFORMATION: NT HIT: U86281.1, EVALUE 0.00e+00
US-09-864-761-30527

  Query Match             52.8%;  Score 535.2;  DB 10;  Length 933;
  Best Local Similarity   74.8%;  Pred. No. 3.1e-164;
  Matches  697;  Conservative    0;  Mismatches  233;  Indels    2;  Gaps    2;

Qy      52 GATATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGGC 111
           || |||||||| |||   |||    |||||||||||||||||||||||||| | | |||
Db     933 GAAATGGGGGAAAATCAGACAATGGTCACAGAGTTCCTCCTACTGGGATTTCTCCTGGGC 874

Qy     112 CCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTACGTCTTCACCCTGCTG 171
           ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||
Db     873 CCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCTATATCTTCACCCTGCTG 814

Qy     172 GGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTAC 230
           |||||||| |||| |||||||||||||||||||||||||||||| ||| ||||| |||||
Db     813 GGGAACGGGGCCATCCTGGGGCTCATCTCACTGGACTCCAGACTCCACACCCCCATGTAC 754

Qy     231 TTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGG 290
           |||||||||||||||||| |||||||||||||||| || ||||||||||||||| |
Db     753 TTCTTCCTCTCACACCTGGCTGTCGTCGACATCGCCTACACCCGCAACACGGTGCCCCAG 694

Qy     291 ATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAG 350
           ||||||| |||||||||||||||||||||||||||||||||||||| ||  |||||| ||||
Db     693 ATGCTGGCGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCTGGTTGCATGACGCAG 634

Qy     351 ACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGATGTCCTAT 410
           |||||||| | ||  | |||||    ||  |||||||||||| |||||| ||||||||||
Db     633 ACCTTTCTCTGTTTGAGTTTTGGACACAGCGAATGTCTCCTGCTGGTGCTGATGTCCTAC 574

Qy     411 GATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCATCATGACCTGGAGAGTC 470
           ||||  |||||||||||||||||||||| ||||||||  |  | ||||||||||||||||||
Db     573 GATCGTTACGTGGCCATCTGCCACCCTCTCCGATACTCCGTCATCATGACCTGGAGAGTC 514

Qy     471 TGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTCCTTTTATCCTTGATTCATCTT 530
           |||||||||||  ||  || ||||||||||    |||   |||  |  |  ||   ||| |
Db     513 TGCATCACCCTGGCCGTCACTTCCTGGACGTGTGGCTCCCTCCTGGCTCTGGCCCATGTG 454

Qy     531 GTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGA 590
           ||  |  | |||    | |||||| | |||| ||  ||||  | ||| || || ||||||
Db     453 GTTCTCATCCTAAGACTGCCCTTCTCTGGGCCTCATGAAATCAACCAC-TTCTTCTGTGA 395

Qy     591 AATCTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTT 650
           |||| || |||| |||  || |||||||| || |||    ||||  ||    | ||||
Db     394 AATCCTGTCTGTCCTCAGGCTGGCCTGTGCTGACACCTGGCTCAACCAGGTGGTCATCTT 335

Qy     651 GGCCGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATATATGTG 710
           || |      | |     |||||||| ||      |    |  | || || ||   |
Db     334 TGCAGCCTGCGTGTTCTTCCTGGTGGGGCCACCCAGCCTGGTGCTTGTCTCCTACTCGCA 275

Qy     711 CATCCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCAC 770
           ||||||    || |||||   |||||| ||  |||| |  |  || || ||||||| |||
Db     274 CATCCTGGCGGCCATCCTGAGGATCCAGTCTGGGGAGGGCCGCAGAAAGGCCTTCTCCAC 215

Qy     771 CTGCTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGGCACAGCCATTATCATGTATGT 830
           ||||| |||||||||||| ||| | |||||||| | |||||  |||||  |||||||||  |
Db     214 CTGCTCCTCCCACCTCTGCGTGGTGGGACTCTTCTTTGGCAGTGCCATCATCATGTACAT 155

Qy     831 TGGACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCT 890
           |  ||||  |   |  | ||  |||||||| | ||     |  | || ||| ||||  |
Db     154 GGCCCCCAAGTCCCGCCATCCTGAGGAGCAGCAAAAGGTCTTTTTTCTATTTTACAGTTT 95

Qy     891 CTTTAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATAC 950
           || || || |  || || ||||| || |  || || ||||||  ||| || |||  | |
Db      94 TTTCAACCCAACACTTAACCCCCTGATTTACAGCCTGAGGAACGGAGAGGTCAAGGGTGC 35

Qy     951 TTTGAAGAGAGTGCTGGGAGTAGAAAGGGCTT 982
           ||| |||||  |||||    |||||   ||
Db      34 CCTGAGGAGAGCACTGGGCAAGGAAAGTCATT 3


RESULT 15
US-09-747-835A-41
; Sequence 41, Application US/09747835A
; Patent No. US20020146692A1
; GENERAL INFORMATION:
;  APPLICANT: Yamazaki, Victoria
;  APPLICANT:  Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Drmanac, Radoje T
;  TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-L
;  TITLE OF INVENTION:  LIKE) POLYPEPTIDES AND POLYNUCLEOTIDES
;  FILE REFERENCE: HYS-37CIP
;  CURRENT APPLICATION NUMBER: US/09/747,835A
;  CURRENT FILING DATE:  2002-03-08
;  PRIOR APPLICATION NUMBER: US 09/729,739
;  PRIOR FILING DATE: 2000-12-04
```

```
;  PRIOR APPLICATION NUMBER: US 09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US 09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US 09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 41
;   LENGTH: 1782
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (485)..(1411)
US-09-747-835A-41

  Query Match           50.4%;  Score 511.4;  DB 10;  Length 1782;
  Best Local Similarity  72.7%;  Pred. No. 2.8e-156;
  Matches  703;  Conservative    0;  Mismatches  256;  Indels    8;  Gaps    3;

Qy      31 CTTGATTTCCTCTGTCATAGGGATATGGGGGACAATATAACATCCATCACAGAGTTCCTC 90
           |||  |||  | | |  ||| || |||| | | |||   |||    ||||||||||||||
Db     461 CTTTGTTTGTTTTTTAGTAGTGAAATGGTGAAAAATCAGACAATGGTCACAGAGTTCCTC 520

Qy      91 CTACTGGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 150
           |||||||||||||| | | ||||||||||||||||||||||||||||||||||||||||
Db     521 CTACTGGGATTTCTCCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTG 580

Qy     151 TTCTACGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTGGACTCC 210
           |||||  ||||||||||||||||||||||| |||| ||||||||||||||||||||||||
Db     581 TTCTATATCTTCACCCTGCTGGGGAACGGGGCCATCCTGGGGCTCATCTCACTGGACTCC 640

Qy     211 AGACTGCACGCCCCC-TGTACTTCTTCCTCTCACACCTGGCGGTCGTCGACATCGCCTAC 269
           |||||  ||  |||||  |||||||||||||||||||||||| ||||||||||||||||||
Db     641 AGACTCCATACCCCCATGTACTTCTTCCTCTCACACCTGGCTGTCGTCGACATCGCCTAC 700

Qy     270 GCCTGCAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCC 329
           ||  ||||||||||||||| ||||||||  ||||||||||||||||||||||||||||||
Db     701 ACCCGCAACACGGTGCCCCAGATGCTGGCGAACCTCCTGCATCCAGCCAAGCCCATCTCC 760

Qy     330 TTTGCGGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTC 389
           |||||  ||  |||||| |||||||||||| | ||  | |||||    ||  |||||||||
Db     761 TTTGCTGGTTGCATGACGCAGACCTTTCTCTGTTTGAGTTTTGGACACAGCGAATGTCTC 820

Qy     390 CTCCTGGTGGTGATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTG 449
           ||  |||||| |||||||||| ||||  |||||||||||||||||||| |||||||| |
Db     821 CTGCTGGTGCTGATGTCCTACGATCGTTACGTGGCCATCTGCCACCCTCTCCGATACTCC 880

Qy     450 GCCATCATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACTGGAGTC 509
           | ||||||||||||      |||||||| || ||  | ||||||||||   |||   |
Db     881 GTCATCATGACCTG------CTGCATCACTCTGGCCATCACTTCCTGGACATGTGGCTCC 934

Qy     510 CTTTTATCCTTGATTCATCTTGTGTTACTTCTACCTTTACCCTTCTGTAGGCCCCAGAAA 569
           ||  |  |  || | ||| |     |  | |||    | ||||| ||| |||| |  ||
Db     935 CTCCTGGCTATGGTCCATGTGAGCCTCATCCTAAGACTGCCCTTTTGTGGGCCTCGTGAA 994

Qy     570 ATTTATCACTTTTTTTTGTGAAATCTTGGCTGTTCTCAAACTTGCCTGTGCAGATACCCA 629
           ||  | ||| || || ||||||||| || |||| ||||  || |||||||| ||||||
Db     995 ATCAACCAC-TTCTTCTGTGAAATCCTGTCTGTCCTCAGGCTGGCCTGTGCTGATACCTG 1053

Qy     630 CATCAATGAGAACATGGTCTTGGCCGGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCAC 689
             ||||  ||    |  |||| || |    || |     ||||||||||  | | |
Db    1054 GCTCAACCAGGTGGTCATCTTTGCAGCCTGCATGTTCATCCTGGTGGGACCACTCTGCCT 1113

Qy     690 AATTGTAGTTTCATATATGTGCATCCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGT 749
             |  | || || ||      ||||||    || |||||   |||||| ||  |||| |
Db    1114 GGTGCTGGTCTCCTACTCACACATCCTGGCGGCCATCCTGAGGATCCAGTCTGGGGAGGG 1173

Qy     750 TCAGAGGAAAGCCTTCTGCACCTGCTTCTCCCACCTCTGTGTGATTGGACTCTTTTATGG 809
           |  || || |||||||| |||||||| |||||||||||| ||  | |||||||| | |||
Db    1174 CCGCAGAAAGGCCTTCTCCACCTGCTCCTCCCACCTCTGCGTAGTGGGACTCTTCTTTGG 1233

Qy     810 CACAGCCATTATCATGTATGTTGGACCCAGATATGGGAACCCCAAGGAGCAGAAGAAATA 869
           ||  |||||  ||||||| | |  || |  |  | || |||||||| ||||
Db    1234 CAGCGCCATCGTCATGTACATGGCCCCTAAGTCCCGCCATCCTGAGGAGCAGCAGAAGGT 1293

Qy     870 TCTCCTGCTGTTTCACAGCCTCTTTAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAG 929
           ||  | || ||| ||||    || || || ||||| || ||||| || | |  || ||
Db    1294 CCTTTTTCTATTTTACAGTTCTTTCAACCCGATGCTAAACCCCCTGATTTACAACCTGAG 1353

Qy     930 GAACTCAGAAGTGAAGAATACTTTGAAGAGAGTGCTGGGAGTAGAAAGGGCTTTATGAAA 989
           |||   ||| || |||  | |  ||| |||||  ||| |    |||||   ||  | | |
Db    1354 GAATGTAGAGGTCAAGGGTGCCCTGAGGAGAGCACTGTGCAAGGAAAGTCATTCCTAAGA 1413

Qy     990 AGGATTA 996
           |  | |
Db    1414 GGTGTGA 1420


Search completed: May  2, 2003, 06:42:32
Job time : 1681 secs
```

```
OM of: US-09-898-586-24 to: Issued_Patents_NA:*   out_format : pfs

Date: Aug 21, 2002  8:19 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=x1h
 -Q=/cgn2_1/USPTO_spool/US09898586/runat_16082002_205735_20570/app_query.fasta_1.373
 -DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
  NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
  USER=US09898586_@CGN1_1_48 -NCPU=6 -ICPU=3 -LONGLOG
  DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-898-586-24
Query length: 310
Database: Issued_Patents_NA:*
Database sequences: 383533
Database length: 122816752
Search time (sec): 45.940000



score_list:
Sequence          Strd Orig    ZScore    EScore Len  ! Documentation  ..
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-467-948A-1 + 1048.50 2174.83  1.9e-113  1713  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-467-947A-1 + 1048.50 2174.83  1.9e-113  1713  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-748-506-5 +  609.00 1254.94  3.3e-62   966  !  S
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-827-291A-1 +  606.00 1245.13  1.1e-61  1290  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-748-506-7 +  592.00 1219.09  3.2e-60   966  !  S
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-748-506-8 +  578.00 1189.56  1.4e-58   966  !  S
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-748-506-6 +  574.00 1181.13  4.2e-58   966  !  S
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-085-371-5 +  543.00 1116.61  1.7e-54   900  !  S
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-748-506-9 +  452.00  923.64  9.3e-44   984  !  S
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-439-313-526 +  384.00  780.51  8.7e-36   963  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-988-876-2 +  382.50  769.63  3.5e-35  1828  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-465-980-1 +  369.00  743.75  9.7e-34  1474  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-053-303-1 +  369.00  743.75  9.7e-34  1474  !
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US95-07093-1 +  369.00  743.75  9.7e-34  1474
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-706-281A-17 +  185.50  361.73  1.8e-12   978  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-097-231-17 +  185.50  361.73  1.8e-12   978  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-671-525B-9 +  182.50  355.44  4.1e-12   975  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-672-109B-9 +  182.50  355.44  4.1e-12   975  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-842-045-9 +  182.50  355.44  4.1e-12   975  !  S
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-842-238-9 +  182.50  355.44  4.1e-12   975  !  S
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-629-335B-9 +  182.50  355.44  4.1e-12   975  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-599-252-84 -  181.00  348.63  9.9e-12  1320  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-436-074-57 -  181.00  348.63  9.9e-12  1320  !
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US96-06352-84 -  181.00  348.63  9.9e-12  1320
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US96-06583-84 -  181.00  348.63  9.9e-12  1320
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-870-511-7 +  175.00  339.34  3.3e-11   999  !  S
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-671-525B-5 +  174.50  337.34  4.2e-11  1080  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-672-109B-5 +  174.50  337.34  4.2e-11  1080  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-842-045-5 +  174.50  337.34  4.2e-11  1080  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-842-238-5 +  174.50  337.34  4.2e-11  1080  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-629-335B-5 +  174.50  337.34  4.2e-11  1080  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-671-525B-7 +  173.00  335.15  5.6e-11   996  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-672-109B-7 +  173.00  335.15  5.6e-11   996  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-842-045-7 +  173.00  335.15  5.6e-11   996  !  S
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-842-238-7 +  173.00  335.15  5.6e-11   996  !  S
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-629-335B-7 +  173.00  335.15  5.6e-11   996  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-870-511-9 +  173.00  335.12  5.6e-11   999  !  S
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-870-511-11 +  173.00  335.12  5.6e-11   999  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-662-560-1 +  173.00  328.92  1.2e-10  1671  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-780-749A-5 +  173.00  328.92  1.2e-10  1671  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-870-511-5 +  173.00  328.92  1.2e-10  1671
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-686-591-3 +  169.00  316.60  6.0e-10  2308
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-970-715-3 +  169.00  316.60  6.0e-10  2308
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-706-281A-15 +  165.00  312.05  1.1e-09  1671
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-097-231-15 +  165.00  312.05  1.1e-09  1671

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-467-948A-1

seq_documentation_block:
; Sequence 1, Application US/08467948A
; Patent No. 5998164
;  GENERAL INFORMATION:
;    APPLICANT:  LI, YI
;    APPLICANT:  CAO, LIANG
;    APPLICANT:  NI, JIAN
;    APPLICANT:  GENTZ, REINER
;    APPLICANT:  BULT, CAROL J.
;    APPLICANT:  SUTTON III, GRANGER G.
;    APPLICANT:  ROSEN, CRAIG A.
;    TITLE OF INVENTION:  Polynucleotides Encoding Human G-Protein
;    TITLE OF INVENTION:  Coupled Receptor GPR2
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;      STREET:  1100 NEW YORK AVE., NW, SUITE 600
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PATENTIN RELEASE #1.0, VERSION #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/467,948A
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/04079
;      FILING DATE:  30-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  STEFFE, ERIC K.
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1488.1140003/EKS/KLM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1713 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  116..1003
US-08-467-948A-1


alignment_scores:
            Quality: 1048.50              Length:    305
              Ratio:  4.017                Gaps:      4
 Percent Similarity:  85.574   Percent Identity: 69.180

alignment_block:
 US-09-898-586-24 x US-08-467-948A-1   ..

 Align seg 1/1  to: US-08-467-948A-1  from: 1  to: 1713

       1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr 17
         |||       |||   |||   :::||||||||||||||||||||||||
```

```
  128 ATGGTGAAAAATCAGACAATGGTCACAGAGTTCCTCCTACTGGGATTTCT 177

   17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34
      :::||||||||||||||||||||||||||||||||||||||||||||||||
  178 CCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCT 227

   34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  228 ATGTCTTCACCCTGCTGGGGAATGGGACCATCCTGGGGCTCATCTCACTG 277

   51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67
      ||||||||||||||||:::||||||||||||||||||||||||||||||
  278 GACTCCAGACTCCACACCCCCATGTACTTCTTCCTCTCACACCTGGCCGT 327

   67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84
      ||||:::||||||||||||||||||||||||||||:::|||||||||||||
  328 CGTCAACATCGCCTATGCCTGCAACACAGTGCCCCAGATGCTGGTGAACC 377

   84 euLeuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100
      ||||||||||||||||||||||||||||||||||    |||
  378 TCCTGCATCCAGCCAAGCCCATCTCCTTTGCTGGTTGCATGACACTAGAC 427

  101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValValMe 117
      ||||||||   :::||||||   ||||||||||||||||||||:::||
  428 TTTCTCTTTTTGAGTTTTGCACATACTGAATGCCTCCTGTTGGTGCTGAT 477

  117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134
      |||||||||   |||||||||||||||||||||||||||||:::    |
  478 GTCCTACGATCGGTACGTGGCCATCTGCCACCCTCTCCGATATTTCATCA 527

  134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150
      ||||||||||:::||||||||||||||||:::::|||||||||||||
  528 TCATGACCTGGAAAGTCTGCATCACTCTGGGCATCACTTCCTGGACATGT 577

  151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProPh 167
      |||   ||||||:::::::::|||:::   |||:::|||   ||||||||
  578 GGCTCCCTCCTGGCTATGGTCCATGTGAGCCTCATCCTAAGACTGCCCTT 627

  167 eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaValL 184
      ||||   |||:::::|||   ||||||||||||||||||||:::|||||
  628 TTGTGGGCCTCGTGAAATCAACCACTTCTTCTGTGAAATCCTGTCTGTCC 677

  184 euLysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
      ||:::||||||||||||||||||   :::|||:::   :::::::::
  678 TCAGGCTGGCCTGTGCTGATACCTGGCTCAACCAGGTGGTCATCTTTGAA 727

  201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217
      :::::::::      ||||||||||||||||      ::::::||||||||
  728 GCCTGCATGTTCATCCTGGTGGGACCACTCTGCCTGGTGCTGGTCTCCTA 777

  217 rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnA 234
      |      ||||||   :::||||||:::||||||||||   |||   :::|
  778 CTCACACATCCTGGGGGGCATCCTGAGGATCCAGTCTGGGGAGGGCCGCA 827

  234 rgLysAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250
      ||||||||||   ||||||   ||||||||||||||:::|||||||||
  828 GAAAGGCCTTCTCCACCTGCTCCTCCCACCTCTGCGTAGTGGGACTCTTC 877

  251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProLy 267
      :::|||:::||||||:::||||||:::::|||:::      :::|||::
  878 TTTGGSAGCGCCATCGTCATGTACATGGCCCCTAAGTCCCGCCATCCTGA 927

  267 sGluGlnLysLysTyrLeu.LeuLeuPheHisSerLeuPheAsnPro... 282
      :||||||:::|||   ||| :::||||||:::|||   |||:::|||
  928 GGAGCAGCAGAAGGTCCTTTTTCTTATTTTACAGTTCCTTTCAACCCCGA 977

  283 MetLeuAsnProLeuIle.CysSerLeuArgAsnSerGlu.ValLysAsn 298
         ||||||||||||||||    :::||||||||||   ||| ||||||:::
  978 TGCTTAAACCCCCTGATTTACAACCCTGAGGAATGTAGAGGGTCAAGGGT 1027

  299 ThrLeuLys 301
      :::|||:::
 1028 GCCCTCCGA 1036
```

```
seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-467-947A-1

seq_documentation_block:
  Sequence 1, Application US/08467947A
  Patent No. 6090575
;  GENERAL INFORMATION:
;    APPLICANT:  LI, YI
;    APPLICANT:  CAO, LIANG
;    APPLICANT:  NI, JIAN
;    APPLICANT:  GENTZ, REINER
;    APPLICANT:  BULT, CAROL J.
;    APPLICANT:  SUTTON III, GRANGER G.
;    APPLICANT:  ROSEN, CRAIG A.
;    TITLE OF INVENTION:  Polynucleotides Encoding Human G-Protein
;    TITLE OF INVENTION:  Coupled Receptor GPR1
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;      STREET:  1100 NEW YORK AVE., NW, SUITE 600
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PATENTIN RELEASE #1.0, VERSION #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/467,947A
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/04079
;      FILING DATE:  30-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  STEFFE, ERIC K.
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1488.1140002/EKS/KLM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1713 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  116..1003
US-08-467-947A-1


alignment_scores:
            Quality: 1048.50              Length:    305
              Ratio:  4.017                Gaps:      4
 Percent Similarity:  85.574   Percent Identity: 69.180

alignment_block:
 US-09-898-586-24 x US-08-467-947A-1   ..

 Align seg 1/1  to: US-08-467-947A-1  from: 1  to: 1713

       1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPheEr 17
         |||       |||   |||   :::||||||||||||||||||||||||
     128 ATGGTGAAAAATCAGACAATGGTCACAGAGTTCCTCCTACTGGGATTTCT 177

      17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34
         :::||||||||||||||||||||||||||||||||||||||||||||||||
     178 CCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCT 227
```